1683

PATENT

#22
M.P.J
2/19/02

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In re Application of:
Susan Lindquist

Serial No.: 09/207,649

Filed: December 8, 1998

For: METHODS FOR IDENTIFYING FACTORS THAT CONTROL THE FOLDING OF AMYLOID PROTEINS OF DIVERSE ORIGIN

Group Art Unit: 1647

Examiner: S. Turner

Atty. Dkt. No.: ARCD:278/GNS

CERTIFICATE OF MAILING
37 C.F.R. § 1.8

I hereby certify that this correspondence is being deposited with the U.S. Postal Service as First Class Mail in an envelope addressed to: Assistant Commissioner for Patents, Washington, D.C. 20231, on the date below:

January 24, 2002
Date

Gina N. Shishima

APPEAL BRIEF

Table of Contents

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In re Application of:
Susan Lindquist

Serial No.: 09/207,649

Filed: December 8, 1998

For: METHODS FOR IDENTIFYING FACTORS THAT CONTROL THE FOLDING OF AMYLOID PROTEINS OF DIVERSE ORIGIN

Group Art Unit: 1647

Examiner: S. Turner

Atty. Dkt. No.: ARCD:278/GNS

APPEAL BRIEF

BOX AF
Assistant Commissioner of Patents
Washington, D.C. 20231

Sir:

Appellant hereby submits an original and two copies of this Appeal Brief to the Board of Patent Appeals and Interferences in response to the final Office Action dated June 19, 2001. The fee for filing this Appeal Brief is $160.00, and is attached hereto.

The Notice of Appeal was filed on September 19, 2001, and received in the Patent and Trademark Office on September 24, 2001, making this Appeal Brief due on November 24, 2001. A petition for a two-month extension of time to respond is included herewith along with the required fees for filing this Appeal Brief. This two-month extension of time will bring the due date to January 24, 2001. Should any other fees be due, or the attached fee be deficient or absent, the Commissioner is authorized to withdraw the appropriate fees from Fulbright &

Jaworski Deposit Account No. 50-1212/10008013/GNS. Please date stamp and return the enclosed postcard to evidence receipt of this document.

I. REAL PARTY IN INTEREST

The real party in interest is the assignee, ARCH Development Corporation.

II. RELATED APPEALS AND INTERFERENCES

There are no interferences or appeals for related cases.

III. STATUS OF THE CLAIMS

Claims 1-36 were originally filed in the application on December 8, 1998. Claims 1-22, the Group I claims, were elected for prosecution in telephonic conference with the Examiner on July 7, 1998. Claims 1, 3, 4, 7, 12, 15, and 16 were amended, claim 37 was added, and claims 2, 5, 6, and 21 were cancelled without prejudice or disclaimer in the Response to Office Action dated August 4, 1999. Claims 22-36 were cancelled without prejudice or disclaimer, claims 1 and 4 were amended, and claims 38-40 were added in the Response to Office Action dated February 1, 2000. Claim 4 was cancelled without prejudice or disclaimer and claims 1, 7, 12, and 14 were amended in Response to Office Action dated October 24, 2000. Claims 1, 3, 7-20, 22, and 37-40 were rejected in the Final Office Action dated June 19, 2001. Claims 38-40 were withdrawn from consideration as being directed to a non-elected invention in the Final Office Action dated June 19, 2001.

Accordingly, claims 1, 3, 7-20, 22, and 37 are pending. Of these, claims 1, 3, 7-20, 22, and 37 are the subject of the present appeal and stand appealed. A copy of the appealed claims is attached as Appendix A to this Brief.

IV. STATUS OF AMENDMENTS

No amendments have been filed subsequent to the final rejection.

V. SUMMARY OF THE INVENTION

The present invention discloses a method of identifying a candidate substance that inhibits the aggregation of a mammalian aggregate-prone amyloid protein, comprising: (a) contacting a yeast cell that expresses a chimeric aggregate-prone amyloid protein comprising a mammalian aggregate-prone amyloid peptide with said candidate substance under conditions effective to allow aggregated amyloid formation; and (b) determining the ability of said candidate substance to inhibit the aggregation of the aggregate-prone amyloid protein. In other embodiments, the mammalian aggregate-prone amyloid protein comprises a PrP or β-amyloid polypeptide. In further embodiments the chimeric protein comprises at least an aggregate forming domain of a mammalian aggregate-prone amyloid protein operably attached to a detectable marker protein. In certain embodiments, the marker protein is green fluorescent protein, luciferase, a drug-resistance marker protein, or a hormone receptor. In particular embodiments, the hormone receptor is a glucocorticoid receptor. In other embodiments, the chimeric protein comprises at least an aggregate forming domain of PrP or β-amyloid. In certain embodiments, the chimeric protein comprises at least about amino acids 1-42 of β-amyloid protein. In other embodiments, the chimeric protein comprises Sup35 in which the N-terminal domain has been replaced by amino acids 1-42 of β-amyloid protein. In further embodiments, any aggregation of the mammalian aggregate-prone amyloid protein is detected by the ability of the aggregated protein to bind Congo Red or is detected by increased protease resistance of the aggregated protein. In other embodiments, the aggregate-prone amyloid protein is labeled. The label may be a radioactive isotope, a fluorophore, or a chromophore. In certain embodiments,

the label is ^{35}S. In particular embodiments, the fluorophore comprises a green fluorescent protein polypeptide. In other embodiments, the yeast cell overexpresses Hsp104. In further embodiments, the aggregated amyloid formation is evidenced by the formation of fibrillary material. Specification at page 5, lines 10-15, 18-19, and 21-30; page 6, lines 1-30; page 7, lines 1 and 8-9.

VI. ISSUES ON APPEAL

A. Whether claims 1, 3, 7-20, and 22 are properly rejected under 35 U.S.C. § 102(b) as being anticipated by Hughes *et al.*?

B. Whether claims 1, 3, 7-20, 22, and 37 are properly rejected under 35 U.S.C. § 102(b) as being anticipated by Cordell *et al.*?

C. Whether claims 1, 2, 7, 12-13, 17-18, and 37 are properly rejected under 35 U.S.C. § 102(e) as being anticipated by Findeis *et al.*?

D. Whether claims 7-11 are properly rejected under 35 U.S.C. § 112, first paragraph as being not enabled?

E. Whether claims 7-11 are properly rejected under 35 U.S.C. § 112, second paragraph as being indefinite for failing to particularly point out and distinctly claim the subject matter which Appellant regards as the invention?

F. Whether claims 7-11 are properly objected to under 37 C.F.R. § 1.75(c) as being of improper dependent form for failing to further limit the subject matter of a previous claim?

VII. GROUPING OF THE CLAIMS

For purposes of this Appeal, the claims do not stand or fall together as set forth in the Argument below. In summary, the Examiner's only stated grounds for rejecting claims 8-11, 14-16, 19, 20, and 22 as being anticipated by Hughes *et al.* and claims 4, 7-14, 16, 20, 22, and 37 as being anticipated by Cordell *et al.* is that these claims depend from rejected base claims. Appellant contends that it is improper to reject claims on the sole basis that they depend from rejected base claims. In any event, every element of these rejected claims are not disclosed either expressly or inherently in the cited prior art references.

VIII. ARGUMENT

A. Rejections Under 35 U.S.C. § 102 Are Improper

1. Standard of anticipation

The novelty of the claimed invention is tested by determining whether or not the claimed invention is anticipated by the prior art as defined in 35 U.S.C. § 102. Anticipation requires that each and every element of the claimed invention be described, either expressly or inherently, in a single prior art reference. *Telemac Cellular Corp. v. Topp Telecom, Inc.*, 247 F.3d 1316, 1327, 58 U.S.P.Q.2d 1545, 1552 (Fed. Cir. 2001); *Verdegaal Bros., Inc. v. Union Oil Co.*, 814 F.2d 628, 631, 2 U.S.P.Q.2d 1051, 1053 (Fed. Cir. 1987). An anticipation analysis requires identifying the elements in the claims, determining their meaning in light of the specification and prosecution history, and identifying the corresponding elements disclosed in the anticipating reference. *Helifix Limited v. Blok-Lok, Ltd.*, 208 F.3d 1339, 1346, 54 U.S.P.Q.2d 1299, 1303 (Fed. Cir. 2000); *Lindemann Maschinenfabrik v. American Hoist & Derrick Co.*, 730 F.2d 1452, 1458, 221 U.S.P.Q. 481, 485 (Fed. Cir. 1984).

2. Hughes *et al.* does not anticipate claims 1, 3, 7-20, and 22

The Final Office Action ("the Action") dated June 19, 2001 rejects claims 1, 3, 7, 12-13, and 17-18 under 35 U.S.C. § 102(b) as being anticipated by Hughes *et al.*, PNAS, 93:2065-70, 1996 ("Hughes *et al.*"). Specifically, the Action contends that Hughes *et al.* clearly measures the interaction of Aβ peptides (β-amyloid peptides) that aggregate, as evidenced in Figure 1 of this reference. The Action reasons that the Appellant has not defined "aggregates" to exclude the interaction of Aβ monomers and that the direct interaction of these monomers constitutes an "aggregation" as depicted in Figure 1. The Action states that the conditions in Hughes *et al.* appear to be sufficient for aggregation and that the Aβ TT mutant in this reference is an identified substance which inhibits aggregation. From this, the Action concludes that Hughes *et*

al. anticipates the claimed invention. Also, the Action rejects claims 8-11, 14-16, 19, 20, and 22 solely on the basis that they depend from one of the rejected base claims above.

Appellant respectfully traverses this rejection. Hughes *et al.* does not anticipate claims 1, 3, 7-20, and 22 of Appellant's presently claimed invention.

As stated above, anticipation requires that each and every element of the claimed invention be described, either expressly or inherently, in a single prior art reference. *Telemac Cellular Corp. v. Topp Telecom, Inc.*, 247 F.3d 1316, 1327, 58 U.S.P.Q.2d 1545, 1552 (Fed. Cir. 2001).

Appellant claims as the present invention "A method of identifying a candidate substance that inhibits the aggregation of a mammalian aggregate-prone amyloid protein, comprising: (a) contacting a yeast cell that expresses a chimeric aggregate-prone amyloid protein comprising a mammalian aggregate-prone amyloid peptide with said candidate substance under conditions effective to allow aggregated amyloid formation; and (b) determining the ability of said candidate substance to inhibit the aggregation of the aggregate-prone amyloid protein. Claim 1.

The Hughes *et al.* reference (Exhibit 1) makes it clear that the authors are not investigating aggregation, but rather the interaction of monomers. In the opening paragraphs of the article, the rationale for the studies in the article are described and stated as:

> Kinetic studies on AB aggregation have demonstrated that amyloid formation is a nucleation-dependent phenomenon, and that lag time precedes aggregation, the length of which may depend on protein concentration. The nucleation even may therefore be the rate-determining step of *in vivo* amyloidosis...Interaction between two monomers, a thermadynamically unfavorable intermolecular interaction, may be a critical step in nucleation.

Hughes *et al.*, page 2065, column 2 (citations omitted). It is evident that monomer formation is hypothesized as part of a nucleation event that ***precedes aggregation***. Hughes *et al.* goes on to state that "[t]he slow and thermodynamically unfavorable interactions between individual

monomers may be the rate-limiting step in aggregation" (page 2070, column 1). A person skilled in the art could only understand this to mean that the interaction of two monomers is necessary ***prior to the aggregation*** of multiple monomers. As evidenced by Hughes *et al.*, a person skilled in the art would not interpret the term "aggregation" to include the interaction between two individual monomers. Furthermore, Appellant's specification does not define the term "aggregation" to include the interaction between two individual monomers. As stated above, an anticipation analysis requires identifying the elements in the claims, determining their meaning in light of the specification and prosecution history, and identifying the corresponding elements disclosed in the anticipating reference. *Helifix Limited v. Blok-Lok, Ltd.*, 208 F.3d 1339, 1346, 54 U.S.P.Q.2d 1299, 1303 (Fed. Cir. 2000). Since a person skilled in the art would not interpret the presently pending claims as encompassing monomer interactions when read in light of Appellant's specification, the Hughes *et al.* reference does not anticipate the present invention.

Furthermore, the Examiner has assumed that an aggregate includes the interaction of two monomers. There is no reason why this assumption is valid. Because the Appellant's specification does not support this definition, the burden rests on the Examiner to provide evidence that the interaction of two monomers constitutes "aggregation." The Examiner has not met this burden. As stated directly above, the very reference the Examiner cites to support this proposition, Hughes *et al.*, specifically distinguishes between the interaction of two monomers and AB ***aggregation*** (page 2065, column 2). Furthermore, in contending that the interaction of AB monomers in Hughes *et al.* constitutes "aggregation," the Examiner refers to Figure 1. Appellant respectfully points out that the authors of Hughes *et al.* admit that Figure 1 is merely a "schematic representation." Hughes *et al.*, at 2066, legend to Figure 1. Furthermore, the occurrence of ***all*** the events in this figure are not supported by any experimental data and does

not constitute evidence that the reference meets the "under conditions effective to allow aggregated amyloid formation." Thus, the Examiner has not met her burden for rejecting the claim as being anticipated by the Hughes *et al.* reference.

Also, another reference cited by the Examiner, Findeis *et al.*, (Exhibit 2) also uses the term "aggregation" to describe the interaction of multiple peptides and ***not between just two peptides***. For example, Findeis *et al.* states that "[t]he term 'aggregation of B amyloid peptides' refers to a process whereby the peptides associate with each other to form a ***multimeric***, largely insoluble complex." Column 8, lines 11-16. These citations provide evidence that a person skilled in the art is well aware that the formation of a multimeric complex includes more than just the interaction between two peptides. Thus, there is no reason why a person of skill in the art would understand that the interaction of two monomers constitutes "aggregation." Since a person skilled in the art does not understand the term "aggregation" to describe the interaction of two monomers and since the Appellant has not defined "aggregation" to include such a description, Hughes *et al.*, which is said to discuss monomer interactions, does not anticipate the presently claimed invention.

Also, Hughes *et al.* describes a typical use of the yeast two-hybrid system. In fact, Figure 1 of this reference precisely shows that the system can evaluate the ability of ***only*** monomers to associate. The steps of Appellant's claimed invention specifically evaluate "aggregated amyloid formation." In fact, the assay in Appellant's claimed (Claim 1) invention is set up "under conditions effective to allow aggregated amyloid formation." In contrast, the two-hybrid system in Hughes *et al.* relies on the ability of two molecules to interact in the nucleus to promote transcription of a reporter gene, which is subsequently assayed. The Hughes *et al.* reference states, "The yeast system described in this paper offers an opportunity to study the interaction of monomeric AB peptides....This system may therefore provide an opportunity to freeze-frame the

monomer-monomer interaction." Hughes *et al.* p. 2070, column 1. Figure 1 in Hughes *et al.* also confirms this by stating: "The system therefore provides an opportunity to examine interaction between two monomeric AB molecules, an essential first step in the nucleation event leading to fibril formation." Hughes *et al.* page 2066. Appellant's specification states that "amyloid or amyloid like deposits are generally insoluble fibrillary material." Specification at page 5, lines 18-19. This indicates not only that the method of Hughes *et al.* is not performed under conditions to promote "aggregated amyloid formation," which is a limitation recited by the claims, but also that it ***would not be*** performed under such conditions because such aggregation of the monomers would prevent transcription—the very event being assayed—from occurring at all.

Furthermore, nowhere in Hughes *et al.* is there a suggestion that the assay be performed under conditions to promote aggregated amyloid formation, which is not surprising given that the assay was not intended to evaluate amyloid formation. The limitation that steps of the assay be performed "under conditions effective to allow aggregated amyloid formation " is conspicuously absent from Hughes *et al.* The yeast two-hybrid system is a widely used, well understood assay that relies on the ability of two polypeptides to bind DNA in a specific (as opposed to non-specific) manner and to associate in a stereo-specific way with proteins that form part of the host cell's transcriptional machinery. This assay is critically dependent on the ability of the polypeptides involved in the assay to reach the nucleus so they can specifically interact with one another to transcribe the reporter gene being assayed. *See* Phizicky and Fields, *Microbiol. Rev.* 59:94-123, 106 (1995) ("The two-hybrid system is limited to proteins that can be localized to the nucleus, which may prevents its use with certain extracellular proteins.") (Exhibit 3). In the assays of the present invention, aggregation prevents the polypeptides from reaching the nucleus. In the context of the Hughes *et al.* reference, aggregation of a mammalian aggregate-prone

amyloid proteins causes them to be insoluble such that the A monomers in the two-hybrid system would be unable to associate in the nucleus, and they would be unavailable to promote transcription. Therefore, the yeast-two hybrid system is simply inoperable for the intended purpose of the assay if practiced according to the limitations of the claimed invention. Thus, Hughes *et al.* does not suggest that the assay be performed under conditions to promote aggregated amyloid formation.

Since the Hughes *et al.* reference does not teach or suggest all of the claim limitations in Appellant's claimed invention, it does not anticipate the presently claimed invention.

Accordingly, Appellant respectfully requests that the rejection of claims 1, 3, 7-20, and 22 under 35 U.S.C. § 102(b) as being anticipated by Hughes *et al.* be withdrawn.

3. Claims 8-11, 14-16, 19, 20, and 22 do not stand or fall with the rejected claims 1, 3, 7, 12-13, and 17-18 as being anticipated by Hughes *et al.*

The Action rejects claims 1, 3, 7, 12-13, and 17-18 under 35 U.S.C. § 102(b) as being anticipated by Hughes *et al.* The Action then proceeds to reject claims 8-11, 14-16, 19, 20, and 22 on the grounds that these claims depend from the rejected base claims.

Appellant takes the position that claims 8-11, 14-16, 19, 20, and 22 are separately patentable from rejected base claims 1, 3, 7, 12-13, and 17-18.

Appellant respectfully points out that it is improper to reject claims on the sole basis that they depend from a rejected base claim. In fact, the *Manual of Patent Examining Procedure* states that "[i]f the base claim is rejected, the dependent claim should be objected to rather than rejected, if it is otherwise allowable." *Manual of Patent Examining Procedure* § 608.01(n)(v). Furthermore, since the Hughes *et al.* reference does not teach or suggest the limitations claimed in claims 8-11, 14-16, 19, 20, and 22 as discussed in the following paragraphs, these claims are

otherwise allowable. Therefore, it is improper to reject 8-11, 14-16, 19, 20, and 22 on the sole basis that they depend from rejected base claims.

As mentioned above, the Hughes *et al.* reference does not teach or suggest every element of claims 8-11, 14-16, 19, 20, and 22 either expressly or inherently. Specifically, claim 8 discloses a chimeric protein comprising at least an aggregate forming domain of a mammalian aggregate-prone amyloid protein operably attached to a detectable marker protein, wherein the marker protein is green fluorescent protein or luciferase. Claim 9 discloses that the marker protein is a drug-resistance marker protein. Claim 10 states that the marker protein is a hormone receptor and claim 11 further states that the hormone receptor is a glucocorticoid receptor. Hughes *et al.* does not teach or suggest any of the above mentioned elements claimed in the present invention. In fact, the Action implicitly concedes this by stating that the only basis for rejection of claims 8-11 is that they depend from a rejected claim.

Also, claim 14 claims "The method of claim 1, wherein the chimeric protein comprises Sup35 in which the N-terminal domain has been replaced by amino acids 1-42 of β-amyloid protein." Claim 15 states "The method of claim 1, wherein any aggregation of the mammalian aggregate-prone amyloid protein is detected by the ability of the aggregated protein to bind Congo Red." Claim 16 further states "The method of claim 1, wherein any aggregation of the mammalian aggregate-prone amyloid protein is detected by increased protease resistance of the aggregated protein." Hughes *et al.* does not teach or suggest any of the above mentioned elements claimed in claims 14-16. In fact, the Action concedes this by stating that the only basis for rejection of claims 14-16 is that they depend from a rejected claim.

Claim 19 claims the method of claim 1, wherein the aggregate-prone amyloid protein is labeled, wherein the label is a radioactive isotope, a flourophore, or a chromophore, wherein the label is ^{35}S. Claim 20 states "The method of claim 18, wherein the flourophore comprises a

green flourescent protein polypeptide." Claim 22 claims "The method of claim 1, wherein said yeast cell over expresses Hsp104." Again, Hughes *et al.* does not teach or suggest any of the above mentioned elements claimed in claims 19, 20, or 22. In fact, the Action concedes this by stating that the only basis for rejection of claims 19, 20, and 22 is that they depend from rejected claims.

Since it was improper to reject claims 8-11, 14-16, 19, 20, and 22 on the sole basis that these claims depend from rejected base claims 1, 3, 7, 12-13, and 17-18 and since claims 8-11, 14-16, 19, 20, and 22 are separately patentable from the rejected base claims, claims 8-11, 14-16, 19, 20, and 22 do not stand or fall with rejected claims 1, 3, 7, 12-13, and 17-18. Furthermore, as discussed above, the Hughes *et al.* reference does not teach or suggest every element of claims 8-11, 14-16, 19, 20, and 22 either expressly or inherently.

Accordingly, Appellant respectfully requests that the rejection of claims 8-11, 14-16, 19, 20, and 22 be withdrawn.

4. Cordell *et al.* does not anticipate claims 1, 3, 15, 17-19, and 37

The Action rejects claims 1, 3, 15, 17-19, and 37 under 35 U.S.C. § 102(b) as being anticipated by Cordell *et al.*, WO 91/04339. The Action further rejects claims 4, 7-14, 16, 20, 22 and 37 solely on the basis that they depend from a rejected base claim. The Action contends that Appellant's definition of "chimeric" appears to be that of non-naturally occurring. The Action further contends that Cordell *et al.* teaches substitution of one or more amino acids in the resulting amyloid peptides thus constituting non-naturally occurring peptides which are chimeric peptides as the sequences differ from the naturally occurring sequence.

Appellant respectfully traverses this rejection. Cordell *et al.* does not anticipate claims 1, 3, 7-20, 22, and 37 of Appellant's presently claimed invention.

As stated above, anticipation requires that each and every element of the claimed invention be described, either expressly or inherently, in a single prior art reference. Cordell *et al.* (Exhibit 4) discloses a single amyloid polypeptide with substitutions of one or more amino acids. Cordell *et al.* does not teach a "chimeric aggregate prone amyloid protein" anywhere in the disclosure which is a required element of claim 1 in the present invention. Furthermore, there is no suggestion to employ the use of a chimeric protein in Cordell *et al.* The Action's contention that the substitution of one or more amino acids in a single amyloid peptide constitutes a chimeric protein is incorrect. It is well known to people skilled in the art that a chimeric protein is the fusion of two separate proteins into one whole protein. In fact, Appellant's specification defines "chimeric protein" to mean "the protein comprises polypeptides that do not naturally occur together in a single protein unit." Specification at page 5, lines 26-27. Appellant's specification uses the term polypeptides in the plural form. This shows an explicit intent to define a "chimeric protein" as including ***at least two*** separate polypeptides that do not naturally occur together. In stark contrast, the Action contends that a ***single*** polypeptide with a substituted amino acid is considered a "chimeric protein." When Appellant's claims are read in light of the specification, a "chimeric protein" comprises of ***at least two*** separate polypeptides that do not naturally occur together. Since Cordell *et al.* does not teach or suggest a "chimeric aggregate-prone amyloid protein" as defined by the specification, it does not teach every element of the presently claimed invention. Therefore, Cordell *et al.* does not anticipate claims 1, 3, 7-20, 22, and 37.

Accordingly, Appellant respectfully requests that the rejection of claims 1, 3, 7-20, 22, and 37 under 35 U.S.C. § 102(b) as being anticipated by Cordell *et al.* be withdrawn.

5. Claims 7-14, 16, 20, 22 and 37 do not stand or fall with the rejected claims 1, 3, 15, 17-19, and 37 as being anticipated by Cordell *et al.*

The Action rejects claims 1, 3, 15, 17-19, and 37 under 35 U.S.C. § 102(b) as being anticipated by Cordell *et al.* The Action then proceeds to reject claims 7-14, 16, 20, 22, and 37 on the grounds that these claims depend from the rejected base claims.

Furthermore, it is unclear from the Action whether the Examiner rejects claim 37 as being anticipated by Cordell *et al.* or as depending from the rejected base claims. In any event, Appellant contends that claim 37 along with claims 7-14, 16, 20, and 22 are separately patentable from rejected base claims 1, 3, 15, and 17-19.

As stated above, it is improper to reject claims on the sole basis that they depend from a rejected base claim. *Manual of Patent Examining Procedure* § 608.01(n)(v). Furthermore, since the Cordell *et al.* reference does not teach or suggest the limitations claimed in claims 7-14, 16, 20, 22, and 37 as discussed in the following paragraphs, these claims are otherwise allowable. Therefore, it is improper to reject 7-14, 16, 20, 22, and 37 on the sole basis that they depend from rejected base claims.

As mentioned above, the Cordell *et al.* reference does not teach or suggest every element of claims 7-14, 16, 20, 22, and 37 either expressly or inherently. Specifically, claim 7 states "The method of claim 1, wherein the chimeric protein comprises at least an aggregate forming domain of a mammalian aggregate-prone amyloid protein operably attached to a detectable marker protein." Claims 8-11 all depend from claim 7. Thus, the limitation in claim 7 extends to claims 8-11. Cordell *et al.* discloses a single amyloid polypeptide with substitutions of one or more amino acids. Cordell *et al.* does not teach or suggest "...a mammalian aggregate-prone amyloid protein operably attached to a detectable marker protein." In fact, the Action concedes this by stating that the sole basis for rejecting claims 7-11 is that they depend from a rejected base claim.

Claim 12 states "The method of claim 1, wherein the chimeric protein comprises at least an aggregate forming domain of PrP or β-amyloid." Claim 13 depends from claim 12. Thus the limitation in claim 12 extends to claim 13. Cordell *et al.* does not teach or suggest "...an aggregate forming domain of PrP or β-amyloid." In fact, the Action concedes this by stating that the sole reason for rejecting claims 12-13 is that they depend from a rejected base claim.

Claim 16 states "The method of claim 1, wherein any aggregation of the mammalian aggregate-prone amyloid protein is detected by increased protease resistance of the aggregated protein." Claim 20 states "The method of claim 18, wherein the flourophore comprises a green fluorescent protein polypeptide" and claim 22 claims "The method of claim 1, wherein said yeast cell over expresses Hsp 104." Finally, claim 37 claims "The method of claim 1, wherein aggregated amyloid formation is evidenced by the formation of fibrillary material." Cordell *et al.* does not teach or suggest any of the limitations claimed in claims 16, 20, 22, or 37. In fact, the Action concedes this by stating that the sole basis for rejecting these claims is that they depend from rejected base claims.

Since it was improper to reject claims 7-14, 16, 20, 22, and 37 on the sole basis that these claims depend from rejected base claims 1, 3, 15, and 17-19 and since claims 7-14, 16, 20, 22, and 37 are separately patentable from the rejected base claims, claims 7-14, 16, 20, 22, and 37 do not stand or fall with rejected claims 1, 3, 15, and 17-19. Furthermore, as discussed above, the Cordell *et al.* reference does not teach or suggest every element of claims 7-14, 16, 20, 22, and 37 either expressly or inherently.

Accordingly, Appellant respectfully requests that the rejection of claims 7-14, 16, 20, 22, and 37 be withdrawn.

6. Findeis *et al.* does not anticipate claims 1, 3, 7, 12-13, 17-18 and 37

The Action rejects claims 1, 3, 7, 12-13, 17-18, and 37 under 35 U.S.C. § 102(e) as being anticipated by U.S. Patent No. 5,854,204 to Findeis *et al.* The Action contends that Findeis *et al.* teaches A-beta peptides including chimeric peptides, which differ from naturally occurring beta amyloid sequences and screening assays to identify modulatory influences on amyloid aggregation.

Appellant respectfully traverses this rejection. Findeis *et al.* does not anticipate claims 1, 3, 7, 12-13, 17-18, and 37.

Appellant claims "A method of identifying a candidate substance that inhibits the aggregation of a mammalian aggregate-prone amyloid protein, comprising: (a) contacting a yeast cell that expresses a chimeric aggregate-prone amyloid protein comprising a mammalian aggregate-prone amyloid peptide with said candidate substance under conditions effective to allow aggregated amyloid formation; and (b) determining the ability of said candidate substance to inhibit the aggregation of the aggregate-prone amyloid protein." Claim 1.

In contrast, Findeis *et al.* (Exhibit 2) never teaches or suggests performing the screening assay in yeast. This is evident in light of Example 5 in the Findeis *et al.* reference (column 51, lines 25 to column 53, line 27). Example 5 in Findeis *et al.* employs the use of a seeded static assay and a shaken plate aggregation assay to identify B-Amyloid modulators. Findeis *et al.* discloses that these assays simply mix a candidate modulator with AB-monomers to determine the affects of the candidate modulator. In fact, these screening assays employed by the Findeis *et al.* reference do not employ the use of yeast cells. In stark contrast, Appellant's screening assay comprises "***contacting a yeast cell*** that expresses a chimeric aggregate-prone amyloid protein comprising a mammalian aggregate-prone amyloid peptide with said candidate substance under

conditions effective to allow aggregated amyloid formation." Thus, Findeis *et al.* does not teach or suggest every element claimed in Appellant's presently claimed invention.

Furthermore, Findeis *et al.* fails to teach or suggest ***incubating*** yeast cells "under conditions effective to allow aggregated amyloid formation," an element expressly required in Appellant's claimed invention. As stated directly above, the screening assays used to identify β-amyloid modulators do not employ the use of yeast cells. Thus, Findeis *et al.* fails to teach or suggest every element required in Appellant's claimed invention either expressly or inherently. Thus, Findeis *et al.* does not anticipate claims 1, 3, 7, 12-13, 17-18, and 37.

Furthermore, Findeis *et al.* does not render claims 1, 3, 7, 12-13, 17-18, and 37 obvious. This is evident in light of the disclosure in Example 5 in Findeis *et al.* Example 5 of Findeis *et al.* teaches away from using a yeast cell to perform screening assay for inhibitors of B-Amyloid aggregation. As discussed above, Example 5 in Findeis *et al.* employs the use of screening assays that simply mix a candidate modulator with AB-monomers to determine the effects of the candidate modulator. Also stated above, these assays employed by Findeis *et al.* do not employ the use of yeast cells. Since the assays employed by Findeis *et al.* teach away from employing the use of yeast cells, Findeis *et al.* does not render claims 1, 3, 7, 12-13, 17-18, and 37 obvious.

Accordingly, Appellant respectfully requests that the rejection of claims 1, 3, 7, 12-13, 17-18, and 37 under 35 U.S.C. § 102(e) as being anticipated by Findeis *et al.* be withdrawn.

B. Rejections Under 35 U.S.C. § 112 Are Improper

1. Claims 7-11 are enabled under 35 U.S.C. § 112, first paragraph

The Action rejects claims 7-11 under 35 U.S.C. § 112, first paragraph, as not providing enablement for a mammalian aggregate-prone amyloid protein wherein the protein is a chimeric that comprises an aggregate forming domain. The Action cites Skolnick *et al.*, Trends in

Biotech., 18(1):34-39, 2000 as recognizing that proteins are highly dependent upon sequence structure and that a single mutation of a protein can affect the biological activities of the molecule. The Action contends that the specification fails to teach the structural and functional characteristics of such sequences which direct aggregate formation. The Action further states that specification fails to teach the experimental methodology for such diverse sequences whereby one of skill in the art can readily determine those sequences which are required. The Action also contends that the specification defines no mammalian chimeras. It also alleges that it would take undue trials and errors to practice the claimed invention given the quantity of necessary experimentation, the lack of working examples, the unpredictability of the art, the lack of sufficient guidance and the breadth of the claims.

Appellant respectfully traverses this rejection. Claims 7-11 are enabled by the present specification.

The claims were previously amended to clarify that the claimed methods involve a "yeast cell that expresses a chimeric aggregate-prone amyloid protein comprising a mammalian aggregate-prone amyloid peptide." The amendment reflects that an "aggregate-prone amyloid protein" contains the amino acid sequence of a "mammalian aggregate-prone amyloid peptide," such as all or part of the amyloid protein. Specification at page 5, line 26-page 6, line 6. The term "mammalian" simply refers to the protein sequence source, which is a mammal. The protein could comprise amino acid sequences from two different mammalian sources, which would render it chimeric, according to the disclosure. Specification page 5, lines 28-29 ("By 'chimeric protein' it is meant that the protein comprises polypeptides that do not naturally occur together in a single protein unit.").

In contrast to the statements in the Action, the specification does disclose mammalian chimeras. The specification describes mammalian aggregate-prone amyloid proteins as

including PrP and β-amyloid polypeptide. Specification page 5, lines 22-24. As mentioned above, it also defines a "chimeric protein" as "...the [chimeric] protein comprises polypeptides that do not naturally occur together in a single protein unit." Specification page 5, lines 28-29. Moreover, the specification also describes examples of chimeric proteins that comprise "at least an aggregate forming domain of a mammalian amyloid polypeptide, such as at least amino acids 1-42 of the β-amyloid protein or at least the aggregate forming domain of PrP." Specification page 6, lines 2-4. The disclosure further mentions that the methods of the invention may employ a "protein comprising the aggregate forming domain of the etiological agent of a particular disease in the yeast system to identify therapeutic compounds for that disease," and it lists some amyloidogenic diseases in animals, such as Alzheimer's disease, scrapie, and spongiform encephalopathy. Specification page 7, lines 20-24. The specification states, "Therefore, in determining therapeutic compounds for Alzheimer's disease, one would use a yeast system comprising at least amino acids 1-42 of the β-amyloid protein." Specification page 7, lines 24-26. It also similarly states that the aggregate-forming domain of PrP may be utilized in the methods of the invention. Moreover, a specific example of a chimeric protein comprising the β-amyloid peptide (1-42) and the Sup35 C terminal domain is mentioned as being particularly useful in screening for the [PSI+] phenotype. Specification page 13, lines 22-25. Thus, the disclosure does define mammalian chimeras.

As for the Action's contention that it would require undue trials and errors to practice the claimed invention, Appellant points to the state of the art at the time the application was filed coupled with the disclosure of the application. The reference of Hughes *et al.*, which is cited by the Action, involves fusion proteins. Fusion proteins are chimeric proteins, and the Hughes *et al.* paper describes a fusion protein that contains a portion of the amyloid peptide. It shows how fusion proteins are made and that they are functional. This paper unequivocally supports the

contention that one of skill in the art could make and use chimeric proteins comprising a mammalian aggregate-prone amyloid protein at the time the present patent application was filed, and that the invention can be practiced without undue experimentation.

Furthermore, "The specification must teach those skilled in the art how to make and use the full scope of the claimed invention without 'undue experimentation.'" *Manual of Patent Examining Procedure* § 2164.08 (citing *In re Wright*, 999 F.2d 1557, 1561, 27 U.S.P.Q. 1510, 1513 (Fed. Cir. 1993)). The methods of making such a chimera and employing it in the described methods stated in the specification would not constitute undue experimentation. For example, the specification teaches how to use peptide regions that would be comprised in a chimeric protein. Furthermore, the recombinant techniques for making chimeras and administering them to a yeast cell were well-known at the time the application was filed, as is evidenced by the Hughes *et al.* paper. The *Manual of Patent Examining Procedure* states that "The specification need not disclose what is well-known to those skilled in the art and preferably omits that which is well-known to those skilled and already available in the public." *Manual of Patent Examining Procedure* § 2164.05(a) (citing *inter alia*, *In re Buchner*, 929 F.2d 660, 661, 18 U.S.P.Q. 2d 1331, 1332 (Fed. Cir. 1991)). Thus, the Appellant need not disclose methods for constructing such a chimeric protein because the application discusses various combinations and furthermore, the recombinant technology was readily accessible to the skilled artisan at the time the application was filed.

Also, Appellant respectfully notes that "it is incumbent upon the Patent Office ...to explain *why* it doubts the truth or accuracy of any statement in a supporting disclosure and to back up assertions of its own with acceptable evidence or reasoning which is inconsistent with the contested statement." *Manual of Patent Examining Procedure* § 2164.05 (quoting *In re Marzocchi*, 439 F.2d 220, 224, 169 U.S.P.Q. 367, 370 (CCPA 1971)). The specification

explicitly states that "yeast expressing a chimeric protein comprising the β-amyloid peptide (1-42) and the Sup35 C-terminal domain have a [PSI+] phenotype that leads to cell death." Specification page 13, lines 23-25. Appellant requests an explanation as to why the Examiner doubts this statement and to provide a reference that indicates the claimed invention would not work. The Action's citation of Skolnick *et al.* (Exhibit 5) is irrelevant. The collaborators in Skolnick *et al.* generally discuss an approach to protein function prediction using a sequence-to-structure-function paradigm (page 34, column 2). In fact, this reference is a general discussion of protein function prediction in light of the numerous genes being discovered due to genome-sequencing projects. Thus, Hughes *et al.* has no specific applicability to the presently claimed invention. Even if it did, as discussed above, information in the present specification indicates how to make and use a "chimeric aggregate-prone amyloid protein comprising a mammalian aggregate-prone amyloid peptide."

Thus, Appellant's specification does enable claims 7-11. Furthermore, the arguments provided by the Action do not support a *prima facie* case to the contrary. Accordingly, Appellant respectfully requests that the rejection of claims 7-11 under 35 U.S.C. § 112, first paragraph be withdrawn.

2. Claims 7-11 are definite under 35 U.S.C. § 112, second paragraph

The Action rejects claims 7-11 under 35 U.S.C. § 112, second paragraph, as being indefinite for failing to particularly point out and distinctly claim the subject matter that the Appellant regards as the invention. Specifically, the Action contends that claims 7-11 are indefinite as the skilled artisan is not reasonably apprised of the metes and bounds of "an aggregate forming domain." The Action further states that the skilled artisan has no guidance by which to determine that portion of a chimeric protein which forms "an aggregate forming domain."

Appellant respectfully traverses this rejection. Claims 7-11 are definite and do particularly point out and distinctly claim the subject matter that the Appellant regards as the invention.

The standard for definiteness of a claim is whether a person of skill in the art can determine the scope of the invention based on the language of the claims with "a reasonable degree of certainty." *Manual of Patent Examining Procedure* § 2173.02 (citing *In re Wiggins*, 488 F.2d 538, 179 U.S.P.Q. 421 (C.C.P.A. 1973)).

Claim 7 recites "The method of claim 1, wherein the chimeric protein comprises at least an aggregate forming domain of a mammalian aggregate-prone amyloid protein operably attached to a detectable marker protein." A person of skill in the art could determine that an "aggregate forming domain" refers to the amino acids of an aggregate-prone amyloid protein that are involved in aggregation. In fact, Appellant's specification states that "In an important embodiment, the chimeric protein comprises Sup35 in which the N-terminal domain has been replaced by amino acids 1-42 of β-amyloid protein." This disclosure eliminates any indefiniteness of what constitutes an "aggregate forming domain."

Furthermore, the fact that one skilled in the art may have to figure out what portions of an aggregate-prone amyloid protein constitute an aggregate forming domain does not render the claims 7-11 indefinite. As mentioned above, the *Manual of Patent Examining Procedure* simply requires that a person of skill in the art be able to determine the scope of the claim. In the present case, claims 7-11 employ a phrase that can be easily understood, "aggregate forming domain." Furthermore, the specification uses that phrase consistent with its use in the claims. Finally, the specification identifies ways of determining whether a protein aggregates. In fact, the specification states that "the aggregation may be detected by its ability to bind Congo Red and show apple green birefringence under polarized light." Specification at page 13, lines 2-3.

Thus, not only is the scope of the claims 7-11 are fully ascertainable to a person of ordinary skill in the art and a person of ordinary skill would be able to practice the present invention. Thus, claims 7-11 are not indefinite.

Accordingly, Appellant respectfully requests that the rejection of claims 7-11 under 35 U.S.C. § 112, second paragraph be withdrawn.

C. Objection of Claims 7-11 Is Improper

The Action rejects claims 7-11 under 37 C.F.R. 1.75(c) as being of improper dependent form for failing to further limit the subject matter of a previous claim. The Action contends that the recitation "at least an aggregate forming domain" broadens the scope of claim 1, in particular to the mammalian aggregate-prone amyloid protein.

Appellant respectfully traverses this objection. Claim 7 is written in proper dependent form and further limits the subject matter of claim 1.

Claim 1 states "A method of identifying a candidate substance that inhibits the aggregation of a mammalian aggregate-prone amyloid protein, comprising: (a) contacting a yeast cell that expresses a chimeric aggregate-prone amyloid protein comprising a mammalian aggregate-prone amyloid peptide with said candidate substance under conditions effective to allow aggregated amyloid formation; and (b) determining the ability of said candidate substance to inhibit the aggregation of the aggregate-prone amyloid protein." Claim 1. Claim 7 further limits claim 1 by stating "The method of claim 1, wherein the chimeric protein comprises at least an aggregate forming domain of a mammalian aggregate-prone amyloid protein operably attached to a detectable marker protein." Claim 7. Claim 8 further requires that the "marker protein is green flourescent protein or luciferase. Claim 9 further requires that the "marker protein is a drug-resistance marker protein. Claim 10 further requires that the "marker protein is

a hormone receptor. Claim 11 further requires that the "hormone receptor is a glucorticoid receptor."

It is evident from the claim language that claim 7 further limits claim 1 in two respects. First, claim 7 further requires "an aggregate forming domain of a mammalian aggregate-prone amyloid protein." Second, claim 7 also requires that the chimeric protein is "operably attached to a detectable marker protein." Both of these limitations are not present in claim 1. It is also evident that claims 8-10 further limits claim 7 and that claim 11 further limits claim 10. Thus, since claim 7 further limits claim 1, the objection of claim 7 and its dependent claim 8-11 is improper.

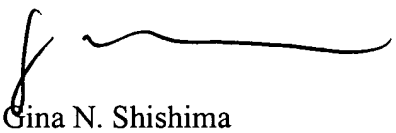
Accordingly, Appellant respectfully requests that the objection of claims 7-11 under 37 C.F.R. 1.75(c) be withdrawn.

IX. CONCLUSION

Appellant has provided arguments that overcome the pending rejections. Appellant respectfully submits that the Office Action's conclusions that the claims should be rejected are unwarranted. It is therefore requested that the Board overturn the Action's rejections.

Please date stamp and return the enclosed postcard to evidence receipt of this document.

Respectfully Submitted,



Gina N. Shishima
Reg. No. 45,104
ATTORNEY FOR APPLICANT

FULBRIGHT & JAWORSKI, L.L.P.
600 Congress Ave, Suite 2400
Austin, Texas 78701
512.536.3055 (voice)
512.536.4598 (fax)

Date: January 24, 2002

APPENDIX A
Pending Claims on Appeal

1. A method of identifying a candidate substance that inhibits the aggregation of a mammalian aggregate-prone amyloid protein, comprising:

(a) contacting a yeast cell that expresses a chimeric aggregate-prone amyloid protein comprising a mammalian aggregate-prone amyloid peptide with said candidate substance under conditions effective to allow aggregated amyloid formation; and

(b) determining the ability of said candidate substance to inhibit the aggregation of the aggregate-prone amyloid protein.

3. The method of claim 1, wherein the mammalian aggregate-prone amyloid protein comprises a PrP or β-amyloid polypeptide.

7. The method of claim 1, wherein the chimeric protein comprises at least an aggregate forming domain of a mammalian aggregate-prone amyloid protein operably attached to a detectable marker protein.

8. The method of claim 7, wherein said marker protein is green fluorescent protein or luciferase.

9. The method of claim 7, wherein said marker protein is a drug-resistance marker protein.

10. The method of claim 7, wherein said marker protein is a hormone receptor.

11. The method of claim 10, wherein said hormone receptor is a glucocorticoid receptor.

12. The method of claim 1, wherein the chimeric protein comprises at least an aggregate forming domain of PrP or β-amyloid.

13. The method of claim 12, wherein the chimeric protein comprises as least about amino acids 1-42 of β-amyloid protein.

14. The method of claim 1, wherein the chimeric protein comprises Sup35 in which the N-terminal domain has been replaced by amino acids 1-42 of β-amyloid protein.

15. The method of claim 1, wherein any aggregation of the mammalian aggregate-prone amyloid protein is detected by the ability of the aggregated protein to bind Congo Red.

16. The method of claim 1, wherein any aggregation of the mammalian aggregate-prone amyloid protein is detected by increased protease resistance of the aggregated protein.

17. The method of claim 1, wherein the aggregate-prone amyloid protein is labeled.

18. The method of claim 17, wherein the label is a radioactive isotope, a fluorophore, or a chromophore.

19. The method of claim 18, wherein the label is ^{35}S.

20. The method of claim 18, wherein the fluorophore comprises a green fluorescent protein polypeptide.

22. The method of claim 1, wherein said yeast cell overexpresses Hsp104.

37. The method of claim 1, wherein aggregated amyloid formation is evidenced by the formation of fibrillary material.

Proc. Natl. Acad. Sci. USA
Vol. 93, pp. 2065–2070, March 1996
Cell Biology

Two-hybrid system as a model to study the interaction of β-amyloid peptide monomers

STEPHEN R. HUGHES, SHEFALI GOYAL, JEANNIE E. SUN, PATRICIA GONZALEZ-DEWHITT, MARYANN FORTES, NORBERT G. RIEDEL, AND SUDHIR R. SAHASRABUDHE*

Neuroscience Therapeutic Domain, Hoechst Marion Roussel Inc., P.O. Box 2500, Somerville, NJ 08876

Communicated by Paul Greengard, The Rockefeller University, New York, NY, November 13, 1995 (received for review July 14, 1995)

ABSTRACT **The kinetics of amyloid fibril formation by β-amyloid peptide (Aβ) are typical of a nucleation-dependent polymerization mechanism. This type of mechanism suggests that the study of the interaction of Aβ with itself can provide some valuable insights into Alzheimer disease amyloidosis. Interaction of Aβ with itself was explored with the yeast two-hybrid system. Fusion proteins were created by linking the Aβ fragment to a LexA DNA-binding domain (bait) and also to a B42 transactivation domain (prey). Protein–protein interactions were measured by expression of these fusion proteins in *Saccharomyces cerevisiae* harboring *lacZ* (β-galactosidase) and *LEU2* (leucine utilization) genes under the control of LexA-dependent operators. This approach suggests that the Aβ molecule is capable of interacting with itself *in vivo* in the yeast cell nucleus. LexA protein fused to the *Drosophila* protein bicoid (LexA-bicoid) failed to interact with the B42 fragment fused to Aβ, indicating that the observed Aβ–Aβ interaction was specific. Specificity was further shown by the finding that no significant interaction was observed in yeast expressing LexA-Aβ bait when the B42 transactivation domain was fused to an Aβ fragment with Phe-Phe at residues 19 and 20 replaced by Thr-Thr (AβTT), a finding that is consistent with *in vitro* observations made by others. Moreover, when a peptide fragment bearing this substitution was mixed with native Aβ-(1–40), it inhibited formation of fibrils *in vitro* as examined by electron microscopy. The findings presented in this paper suggest that the two-hybrid system can be used to study the interaction of Aβ monomers and to define the peptide sequences that may be important in nucleation-dependent aggregation.**

An aggregated form of β-amyloid peptide (Aβ), a 39- to 42-aa peptide, is the principal component of amyloid in the core of plaques, which are characteristic of the Alzheimer disease (AD) brain (1, 2). Aβ is posttranslationally derived from a much larger amyloid precursor protein (APP) encoded by a gene on chromosome 21 in band q21 (3–6). The strongest evidence for involvement of APP in AD comes from familial mutations discovered close to or within the Aβ domain (7–12). At least one of these mutations (codons 670 and 671 of APP-770) has been shown to increase secreted Aβ *in vitro* (13, 14). It is therefore likely that dysfunction in APP or Aβ metabolism may play a role in AD.

Since Aβ was detected in senile plaques, it was assumed that this peptide was a result of abnormal cleavage of APP. It is now accepted that Aβ is secreted by cells in culture and is found as a soluble peptide in the cerebrospinal fluid (CSF) of AD patients and in comparable concentrations in age-matched control patients (15–17). Soluble Aβ has also been detected in the plasma of healthy individuals (15). The measurement of soluble Aβ in CSF of patients free of neurodegenerative disease indicated an increase in peptide levels with age (18). Therefore, physiological factors that can induce Aβ aggregation may be more important in the development of AD pathology than the concentration of Aβ *per se*.

Extracellular and cerebrovascular amyloid deposits are composed of Aβ variants which differ at their carboxyl termini (1, 2, 19, 20). Neuritic plaques have been found to contain high levels of Aβ-(1–42) (1, 19, 21). Kinetic studies on Aβ aggregation have demonstrated that amyloid formation is a nucleation-dependent phenomenon (22), and that a lag time precedes aggregation, the length of which may depend on protein concentration. The nucleation event may therefore be the rate-determining step of *in vivo* amyloidosis.

Nucleation-dependent polymerization is observed in protein crystallization, microtubule assembly, flagellum assembly, phage capsid assembly, actin polymerization (23), and a small subset of human proteins that characterize amyloid diseases (24–26). Interaction between two monomers, a thermodynamically unfavorable intermolecular interaction, may be a critical step in nucleation. The experimental yeast system presented herein provides an opportunity to study the interaction of Aβ monomers *in vivo*.

EXPERIMENTAL PROCEDURES

Bacterial and Yeast Strains. Manipulations of bacterial strains and of DNAs were by standard methods (27, 28) unless otherwise noted. *Escherichia coli* maximum-efficiency DH5α competent cells (GIBCO/BRL) were used as hosts throughout. Yeast strain EGY48 was obtained from the laboratory of Roger Brent (Massachusetts General Hospital, Boston).

Construction of Bait and Prey Plasmids. To construct the bait plasmid (LexA-Aβ fusion), primers to the cDNA for the APP-770 isoform cloned into the *Hin*dIII and *Xba* I sites of pRcCMV were used. For amplification, the primers used were 5′-AAGGCCTGGATCCTGGATGCAGAATTCCGACATGAC-3′ at the 5′ end and 5′-AAGGCCTCTCGAGGTCGACCTACGCTATCAGCAACCACCGCACC-3′ at the 3′ end. This primer set amplified a 163-bp fragment that was digested with *Bam*HI and *Xho* I to obtain a 143-bp *Bam*HI–*Xho* I fragment, which was then ligated into pEG202 (29) at those sites. This places the open reading frame for Aβ into translational phase with the LexA sequence of pEG202. The downstream primer was constructed to contain a synthetic stop codon after Ala-42 of Aβ. The bait fusion protein is produced constitutively from pEG202, a 2-μm *HIS3*$^+$ plasmid under the control of the *ADH1* promoter and encoding the LexA carboxyl-terminal oligomerization region, which contributes to operator occupancy by LexA derivatives (29).

The prey plasmid (B42-Aβ fusion) was constructed by digesting the 163-bp PCR-amplified fragment designed for bait

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviations: Aβ, β-amyloid peptide; AβTT, Aβ peptide with Phe^{19}-Phe^{20} replaced with Thr-Thr; APP, amyloid precursor protein; AD, Alzheimer disease; HA, hemagglutinin; ONPG, *o*-nitrophenyl β-D-galactopyranoside; X-Gal, 5-bromo-4-chloro-3-indolyl β-D-galactopyranoside.

*To whom reprint requests should be addressed.

(see above) with *Eco*RI and *Xho* I. This *Eco*RI–*Xho* I fragment was placed in pJG4-5, a 2-μm *TRP1* plasmid (30, 31), in translational frame with the codons for the simian virus 40 large T nuclear localization signal, the B42 transactivation domain, and the hemagglutinin (HA) epitope tag. The prey fusion protein (16 kDa) will be inducible in yeast grown on minimal medium (MM) containing 2% galactose and 1% raffinose (Gal/Raf) but not in yeast grown on 2% glucose (Glc). Amino acids 3 and 4 (glutamate, phenylalanine) of Aβ (at the point where it is fused to the HA tag) are generated by codons in the *Eco*RI site.

To construct the mutant prey plasmid that contained the AβTT-encoding sequence, a fragment was constructed from a 147-bp oligonucleotide representing the mutation of Phe^{19}-Phe^{20} to Thr^{19}-Thr^{20} within Aβ synthesized on a Millipore model 8909 Expedite nucleic acid synthesis system as follows. The oligonucleotide 5′-AGGCCTGAATTCCGACATGAC-TCAGGATATGAAGTTCATCATCAAAAATTGGT-GACTACTGCAGAAGATGTGGGTTCAAACAAA-GGTGCAATCATTGGACTCATGGTGGGCGGT-GTTGTCATAGCGTAGGTCGACCTCGAGAGGCCT-3′ was annealed with a complementary short oligonucleotide, 5′-AGGCCTCTCGAGGTCGACC-3′, and filled in by Klenow DNA polymerase (BRL). The fragment was extracted with phenol/chloroform and purified with the qiaquick-spin PCR purification kit (Qiagen). The sample was then digested with *Eco*RI and *Xho* I and placed into a ligation reaction mixture in a 7:1 ratio with *Eco*RI/*Xho* I-digested pJG4-5 prey plasmid. The plasmids were propagated and grown in DH5α subcloning-efficiency competent cells from BRL.

The accuracy of the reading frames in the bait and prey plasmids was verified with an automated Applied Biosystems sequencer employing 373 software. Sequences were confirmed to be correct by the analysis features of SEQUENCE EDITOR and MACVECTOR software (data not presented).

Western blot analyses were performed (32) to show that the bait and prey plasmids expressed the expected fusion proteins (data not presented).

Transformation of Strain with Reporter, Bait, and Prey Plasmids. The selection strain was made by transforming the EGY48 yeast strain with a *URA3 lacZ* (β-galactosidase) reporter plasmid and the *HIS3* bait plasmid by the lithium acetate method (27). The yeast selection strain harboring the bait and reporter plasmids was transformed with the prey plasmid DNA (27), and tryptophan utilization phenotype was used (in addition to Ura and His markers for bait and *lacZ* reporter plasmids, respectively) for selection of transformants with prey plasmids.

Determination of Bait–Prey Interaction. Yeast strains containing the appropriate bait and prey plasmids were grown to an OD_{600} of 0.5, diluted 1000-fold, and spotted on plates containing Glc Ura^- His^- Trp^- 5-bromo-4-chloro-3-indolyl β-D-galactopyranoside (X-Gal) medium or Gal/Raf Ura^- His^- Trp^- X-Gal medium to assess the transcriptional activation of the *lacZ* reporter gene. Suitably diluted cell suspensions were also spotted on Gal/Raf Ura^- His^- Trp^- Leu^- medium and Glc Ura^- His^- Trp^- Leu^- medium to assess the transcriptional activation of the *LEU2* gene.

β-Galactosidase Activity in Liquid Cultures of Yeast. Cells were assayed for β-galactosidase activity by the *o*-nitrophenyl β-D-galactopyranoside (ONPG) method (27). The experiment was repeated in triplicate and the plotted data represent an average value of the values for the three samples. The statistical significance was computed with Student's *t* test in a two-tailed analysis.

Immunoprecipitation and Western Blot. Extracts were made from EGY48 cells that contained a prey plasmid encoding B42-Aβ and a bait plasmid encoding LexA-Aβ. Cells were grown in 100 ml of Glc or Gal/Raf medium (in which B42-Aβ expression was induced) to an OD_{600} of 0.6–0.8, pelleted by centrifugation, resuspended in 500 μl of RIPA buffer (25), lysed by beating with glass beads five times for 2 min each, and spun twice for 5 min in a microcentrifuge (10,000 × *g*) at 4°C to remove the beads and cell debris. Five microliters of the supernatant was taken as a control, and 15 μl of rabbit anti-LexA antiserum [kindly donated by Roger Brent (33)] was added to the remainder, which was incubated at 4°C for 4 hr on a rotating platform. LexA-containing proteins were precipitated from this remainder with 50 μl of protein A-Sepharose CL-4B (Sigma). The entire pellet was dissolved in Laemmli sample buffer, subjected to sodium dodecyl sulfate/polyacrylamide gel electrophoresis (SDS/PAGE; Integrated Separation System, Hyde Park, MA), and blotted onto nitrocellulose. Tagged Aβ fusion proteins were identified by Western analysis of the blotted proteins with the 12CA5 monoclonal anti-HA antibody (34). Cell extracts and immunoprecipitates were also subjected to immunoblotting with monoclonal anti-Aβ antibodies 4G8 and 6E10. Western blot analysis was performed with ECL chemiluminescence reagents using the protocol supplied by the vendor (Amersham).

Electron Microscopy. Dilutions of Aβ for incubation with the octapeptide QKLVTTAE were performed as in ref. 41 (ratio of Aβ to octapeptide was 1:10). The photomicrographs were obtained with a JEOL JEM-100S electron microscope at 80 kV (×155,000 magnification).

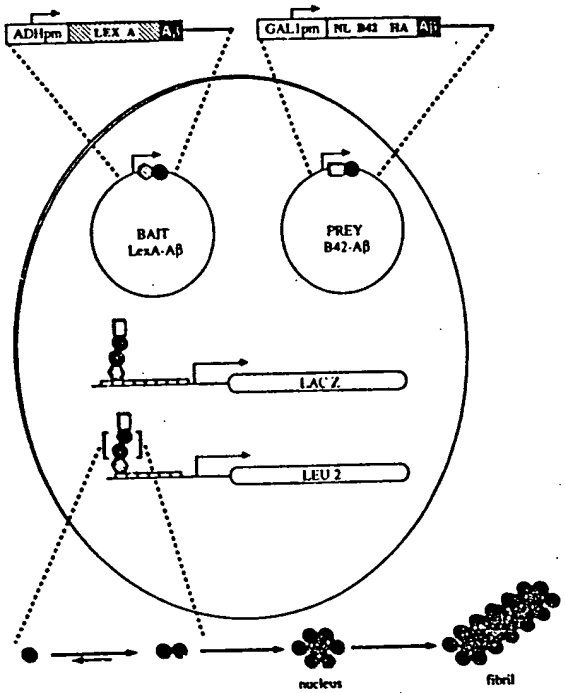


FIG. 1. Schematic representation of yeast strain EGY48 transformed with bait (LexA-Aβ fusion), prey (B42-Aβ fusion), and *lacZ* reporter plasmids. Bait fusion protein (LexA-Aβ) is produced constitutively under the control of the *ADH* promoter and binds to the upstream region of reporter genes at LexA operator sites; prey fusion peptide expression is driven by the *GAL1* promoter and is inducible by galactose. Prey fusion protein is expressed in the presence of galactose. If the Aβ portion of the prey protein binds the Aβ moiety of the bait fusion peptide, transcription from the reporter genes is triggered. The system therefore provides an opportunity to examine interaction between two monomeric Aβ molecules, an essential first step in the nucleation event leading to fibril formation.

RESULTS

The experimental system established by Brent *et al.* (described in ref. 29) is depicted in Fig. 1. The selection strain contains either Aβ or bicoid as a bait fused in-frame to the bacterial LexA protein, which by itself has no transcriptional activation function in yeast (33). The host strain contains *LEU2* and *lacZ* reporters carrying LexA operators instead of native upstream activating sequences. A strain containing the bait (LexA-Aβ) and reporters (*LEU2* and *lacZ*) remains inert for the expression of leucine utilization or β-galactosidase activity unless it also contains a vector (prey) that expresses an interacting protein as a fusion molecule consisting of nuclear localization sequences from simian virus 40, the B42 acid blob transactivation domain, and an epitope tag from influenza virus HA protein (35). In this system, conditional expression of library-encoded proteins is directed by the *GAL1* promoter (achieved by growing yeast cells in Gal/Raf minimal medium).

We first determined whether EGY48 strains containing the LexA protein alone, LexA-Aβ fusion protein, or LexA-bicoid permitted the expression of *lacZ* or leucine genes. When EGY48 strains containing the individual LexA fusion baits were spotted at equal density on minimal medium plates containing Gal/Raf Ura$^-$ His$^-$ medium, similar growth rates were observed, indicating that none of the baits was toxic to yeast. These strains failed to grow on Gal/Raf Ura$^-$ His$^-$ Leu$^-$ medium and did not form blue colonies on Gal/Raf Ura$^-$ His$^-$ X-Gal medium (data not presented), indicating that none of the bait proteins by themselves could permit the expression of leucine or β-galactosidase phenotypes.

B42-Aβ prey plasmid was introduced into the yeast strain containing LexA-Aβ bait protein. Equal dilutions of this yeast strain were spotted on Gal/Raf Ura$^-$ His$^-$ Trp$^-$ X-Gal and Glc Ura$^-$ His$^-$ Trp$^-$ X-Gal media to measure expression of β-galactosidase, and on Gal/Raf Ura$^-$ His$^-$ Trp$^-$ Leu$^-$ and Glc Ura$^-$ His$^-$ Trp$^-$ Leu$^-$ media to check the expression of the leucine utilization phenotype. The B42-Aβ prey plasmid, when introduced into the yeast strain with LexA-Aβ bait, showed growth on minimal medium plates devoid of leucine (Fig. 2*A*) and showed blue colonies on X-Gal medium in the presence of Gal/Raf as the carbon source (Fig. 2*C*) but showed no growth (Fig. 2*B*) and no blue colonies (Fig. 2*D*) in the presence of glucose. These results indicate that the interaction between LexA-Aβ and B42-Aβ is triggered by expression of the B42 fusion protein under the influence of the *GAL1* promoter. When LexA-bicoid (Fig. 2) or LexA protein alone (data not presented) was used as bait, introduction of B42-Aβ plasmid did not result in growth on leucine plates or blue colonies on X-Gal medium in the presence of Gal/Raf, indicating that the

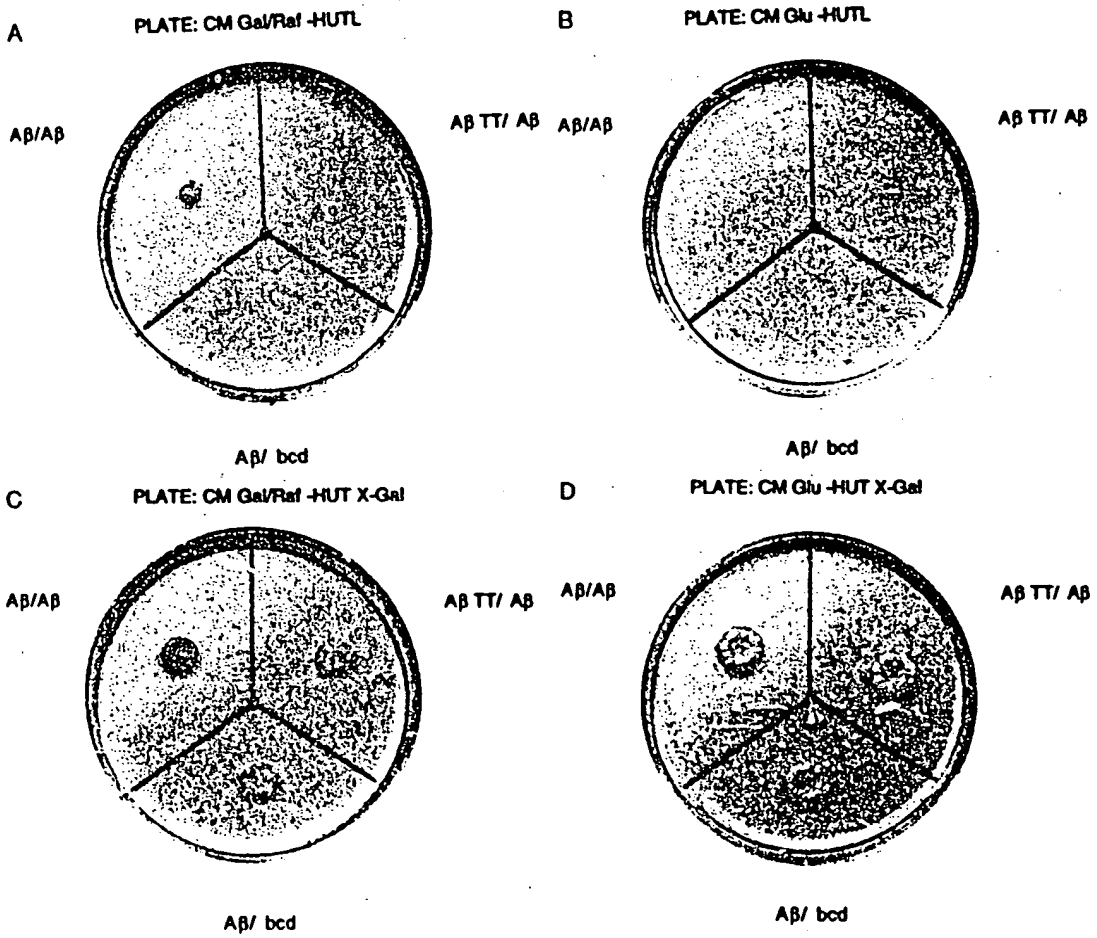


FIG. 2. Four-plate screen to examine interaction. B42-Aβ/LexA-Aβ strain, 342-AβTT/LexA-Aβ strain, and B42-Aβ/LexA-bcd strain were grown on plates containing Gal/Raf Ura$^-$ His$^-$ Trp$^-$ Leu$^-$ minimal medium (*A*), Glc Ura$^-$ His$^-$ Trp$^-$ Leu$^-$ minimal medium (*B*), Gal/Raf Ura$^-$ His$^-$ Trp$^-$ X-Gal minimal medium (*C*), or Glc Ura$^-$ His$^-$ Trp$^-$ X-Gal minimal medium (*D*).

interaction between LexA-Aβ and B42-Aβ is specific and most likely due to the intermolecular interaction between the Aβ molecules derived from the bait and the prey. When B42-AβTT prey plasmid was introduced into the yeast strain bearing LexA-Aβ bait, minimal growth was observed in plates devoid of leucine and no blue (or only very weakly blue) colonies were observed on X-Gal medium in the presence of glucose or Gal/Raf as carbon source (Fig. 2). This indicates that the Aβ molecule substituted at positions 19 and 20 with threonine residues interacts poorly with the wild-type Aβ peptide.

We next attempted to quantitate the observed Aβ–Aβ interaction by the ONPG colorimetric assay. Fig. 3 clearly indicates that there was significantly higher β-galactosidase activity in the yeast strain expressing B42-Aβ prey and LexA-Aβ bait compared with the yeast cells with B42-Aβ prey/LexA-bicoid bait (~2.5-fold, $P = 0.01$, Student's t test) or B42-AβTT prey/LexA-Aβ bait (~2-fold, $P = 0.02$). These results indicate that the Aβ–Aβ interaction inferred from Fig. 2 was significantly greater than the interaction between AβTT and Aβ or Aβ and bicoid monomers.

In an attempt to obtain direct *in vivo* evidence for the interaction between B42-Aβ prey and LexA-Aβ bait proteins, immunoprecipitates obtained by using antibodies against bait protein were subjected to immunoblotting with antibodies against the prey protein. Yeast cells expressing LexA-Aβ bait and B42-Aβ or B42-AβTT prey proteins were grown in glucose-containing medium and switched to glucose or Gal/Raf liquid minimal medium. The cells were harvested and cell extracts were prepared from equal numbers of cells. One aliquot of cell extract was subjected to immunoprecipitation with an anti-LexA antibody and the immunoprecipitates were subjected to Tris/tricine SDS/PAGE followed by immunoblotting with the monoclonal anti-HA antibody. If the two Aβ molecules or Aβ–AβTT molecules interact *in vivo*, one should be able to isolate the bait–prey complexes with antibody specific to the bait. Indeed, prey-specific HA immunoreactivity was observed (at 16 kDa; Fig. 4, lane 1) from the immunoprecipitates obtained from Aβ/Aβ cells grown in the presence of galactose, but not from the immunoprecipitates obtained from these cells subjected to glucose in the medium (Fig. 4, lane 2), indicating that the two Aβ fusion proteins interact inside the yeast cell nucleus. When Aβ/Aβ cell extracts were directly subjected to immunoblotting with anti-HA antibody, the 16-kDa band was observed in extracts derived from these cells grown in the presence of galactose (Fig. 4, lane 3), but no

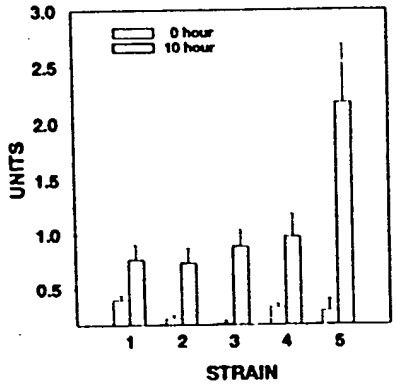


FIG. 3. Assay to determine β-galactosidase activity present in each of the strains tested after 0 hr (open bars) and 10 hr (filled bars) of incubation in Gal/Raf complete minimal medium. Bars: 1, B42-Aβ/LexA-bicoid; 2, B42 alone/LexA alone; 3, B42 alone/LexA-Aβ; 4, B42-AβTT/LexA-Aβ; 5, B42-Aβ/LexA-Aβ. These data are typical of three replicated experiments.

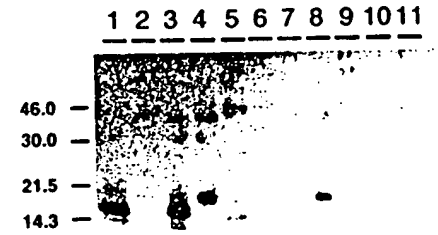


FIG. 4. Immunoprecipitates obtained by using LexA antiserum with extracts of B42-Aβ/LexA-Aβ strain grown in Gal/Raf medium (lane 1) or Glc medium (lane 2), cell extracts of B42-Aβ/LexA-Aβ cells grown in the presence of Gal/Raf (lane 3) or Glc (lane 4). Lanes 5–8 are similar to lanes 1–4 except that B42-AβTT/LexA-Aβ yeast strain was used. Lanes 9 and 10 represent immunoprecipitates and cell extracts obtained from LexA-Aβ strain containing no prey plasmid and from B42 alone/LexA alone yeast strain, respectively. Samples were electrophoresed in a Tris/tricine SDS/15% polyacrylamide gel and immunoblotted with anti-HA antibody.

immunoreactive band was observed for cells grown in the presence of glucose (Fig. 4, lane 4). The LexA immunoprecipitates obtained from AβTT/Aβ yeast grown on galactose resulted in very low levels of the anti-HA-immunoreactive 16-kDa band, seen only upon prolonged exposure (Fig. 4, lane 5; band not seen at this exposure). No 16-kDa band resulted from immunoprecipitates grown on glucose even on prolonged exposure (Fig. 4, lane 6). The anti-HA-immunoreactive 16-kDa band was, however, observed in cell extracts obtained from AβTT/Aβ yeast grown on galactose (Fig. 4, lane 7) but not in AβTT/Aβ yeast grown on glucose (Fig. 4, lane 8; the AβTT prey protein seems to run slightly lower than the Aβ prey protein). The observed molecular mass of 16 kDa is consistent with that predicted for a prey fusion protein and was also observed when anti-Aβ antibodies were used with cell extracts from Aβ/Aβ cells grown on galactose (data not presented). No immunoreactive bands were obtained in LexA immunoprecipitates derived from the EGY48 strain expressing the LexA-Aβ bait but containing no prey plasmid (Fig. 4, lane 9) and in cell extracts or immunoprecipitates from E9Y48 strain.

The observed weak interaction of AβTT with the native Aβ molecule was also examined by electron microscopy. An octapeptide fragment, QKLVTTAE, representing the FF-to-TT substitution at positions 19 and 20 in Aβ-(17–24) is capable of inhibiting fibril formation of the native Aβ-(1–40) (Fig. 5). Aβ-(1–40) at 250 μM, incubated for 4 days in water, showed significantly greater fibril formation (Fig. 5*A*) compared with the amount seen when Aβ-(1–40) at 250 μM was incubated with the octapeptide at 2.5 mM under the same conditions (Fig. 5*B*).

DISCUSSION

In the present study we demonstrate that two monomers of Aβ are capable of interacting in a eukaryotic cell. We further demonstrate that this interaction is specific by using the *Drosophila* protein bicoid as a bait protein in this system. This interaction was found to be positive by using the *lacZ* and *LEU2* reporter systems (Fig. 2). Quantitation of β-galactosidase activity by the ONPG assay supported these conclusions (Fig. 3). Furthermore, direct evidence of interaction was obtained by subjecting immunoprecipitates obtained by using antibodies against bait protein to immunoblotting with antibodies raised against the HA epitope present on the prey protein (Fig. 4).

Hilbich *et al.* (36) have previously reported that a well-preserved hydrophobic core around aa 17–24 is important for

the formation of β-sheet structure and amyloid properties. When stained with Congo red, peptide Aβ-(10–42) or Aβ-(10–43) containing the FF-to-TT substitution (equivalent to the substitution in AβTT) did not reveal birefringence and showed decreased β-sheet content when compared with the native peptide by circular dichroism spectroscopy and by infrared spectroscopy. Moreover, when these substituted peptides were mixed with the native Aβ-(10–43) fragment at equimolar concentration, they inhibited the formation of filaments *in vitro* (below 15%), as detected by electron microscopic analysis (36). Our data suggest that the octapeptide fragment QKLVTTAE, representing the FF-to-TT substitution in Aβ-(17–24), is also capable of inhibiting the aggregation of native Aβ-(1–40) (Fig. 5; unpublished results). The results presented in this paper clearly demonstrate that AβTT fusion protein interacts poorly with the native Aβ (Figs. 2–4). Hilbich *et al.* (36) suggest that the Aβ and AβTT monomers may form oligomers that do not fit into the structure of a filament. Our data suggest that inhibition of filament formation by peptides representing the TT substitution (at 10-fold molar abundance compared with the native peptide) may be explained by a weak interaction between the Aβ and AβTT monomers. Alternatively, it is also possible that AβTT peptide may directly interfere with the fiber formation process.

Recent evidence has indicated that the cellular forms of prion protein (PrP^c) can form protease-resistant prion protein (PrP^{sc}) in a cell-free system in which PrP^{sc} is used as a seed (37). This conversion did not require biosynthesis of new PrP^c, asparagine-linked glycosylation, or the presence of its normal glycosylphosphatidylinositol anchor, suggesting that oligomer

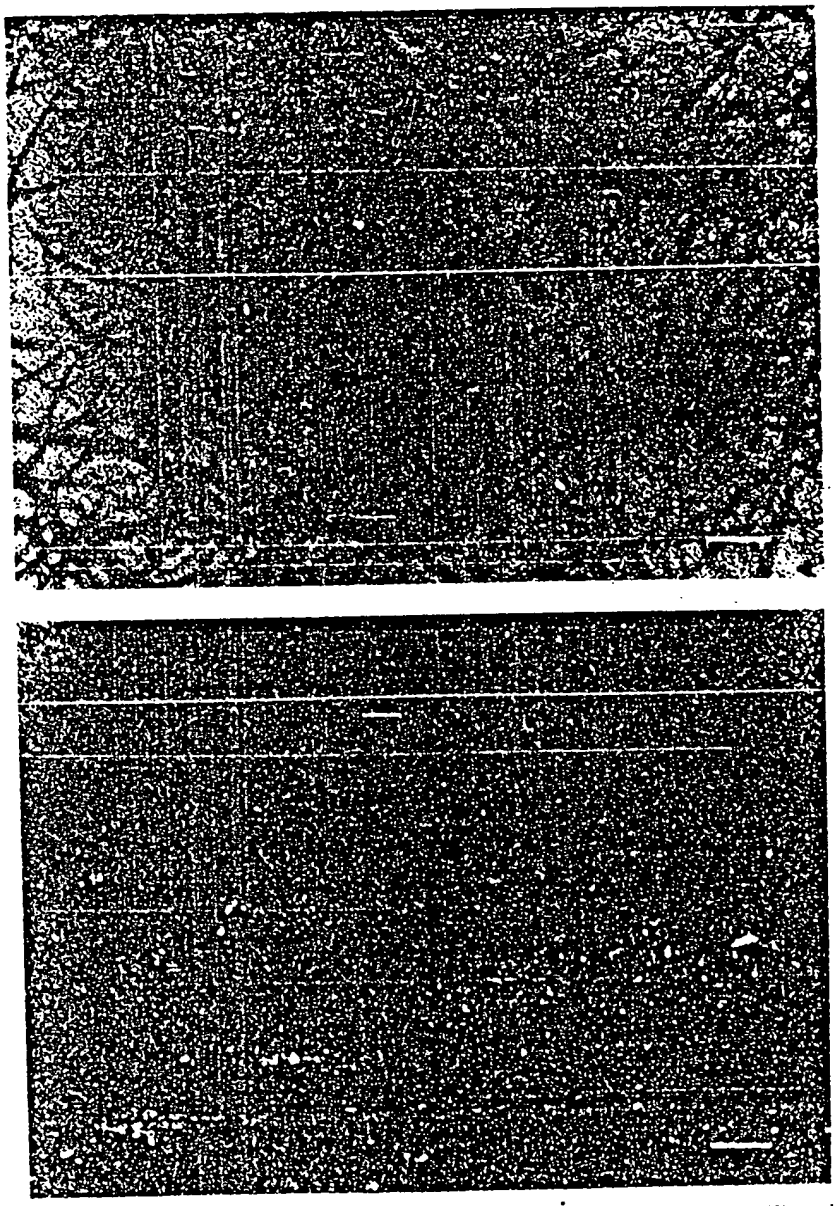


FIG. 5. Transmission electron micrographs of Aβ-(1–40) peptide incubated for 4 days in water (*A*) and Aβ-(1–40) peptide incubated in a 1:10 ratio with the octapeptide QKLVTTAE for 4 days in water (*B*). (Bar = 0.1 μm.)

formation between PrP^{sc} and PrP^{c} is sufficient for the conversion reaction to occur. Amyloid fibrils characterize several human diseases in which the presence of amyloid deposits is coincident with organ dysfunction. There is often a positive correlation between severity of the disease and the extent of fibril formation (38). Amyloid formation exhibits nucleation-dependent kinetics (22, 39–41). The slow and thermodynamically unfavorable interactions between individual monomers may be the rate-limiting step in aggregation. The yeast system described in this paper offers an opportunity to study the interaction of monomeric Aβ peptides. Although the peptides were expressed as fusion proteins, the results presented in this paper suggest that the observed interaction is due to the Aβ peptide domain. Results presented in Fig. 4 also suggest that no covalent higher-order bait–prey aggregates can be observed on the gel. This system may therefore provide an opportunity to freeze-frame the monomer–monomer interaction. Experimental system(s) such as the one presented in this paper may thus be used to study monomer–monomer interactions in other proteins that participate in nucleation-dependent amyloid formation.

There are some caveats in the present study. Although our results indicate that the interaction of LexA-Aβ and B42-Aβ fusion proteins occurs mainly because of amino acids within the two Aβ domains, it is possible that the conformation of the fusion proteins may influence this effect. These interactions take place in the yeast cell nucleus under conditions where the LexA-Aβ bait complex is bound to the LexA operator site. Whereas the yeast system described here is useful in studying intermolecular interactions, the intramolecular interactions may not be fully captured in a fusion-protein context. Furthermore, it is possible that interaction between bait-Aβ and prey-Aβ may not be relevant to fibril formation. This system will therefore have to be carefully validated by using molecules that are known to accelerate or inhibit the monomer–monomer interaction in Aβ fibrillogenesis.

We thank Dr. Roger Brent (Massachusetts General Hospital, Boston) for providing the necessary vectors and yeast strains required for completion of this study. We are very thankful to Dr. Timm Jessen for help with several experimental protocols, Paula Williams for electron microscopy, and Michael Merriman for oligonucleotide and peptide synthesis. The monoclonal antibodies to Aβ peptide were provided by Drs. K. S. Kim and H. Wisniewski.

1. Masters, C. L., Simms, G., Weinman, N. A., Multhaup, G., McDonald, B. L. & Beyreuther, K. (1985) *Proc. Natl. Acad. Sci. USA* **82**, 4245–4249.
2. Glenner, G. G. & Wong, C. W. (1984) *Biochem. Biophys. Res. Commun.* **122**, 1131–1135.
3. Kang, J., Lemaire, H. G., Unterbeck, A., Salbaum, J. M., Masters, C. L., Grzeschik, K. H., Multhaup, G., Beyreuther, K. & Muller-Hill, B. (1987) *Nature (London)* **325**, 733–736.
4. Tanzi, R. E., Gusella, J. F., Watkins, P. C., Bruns, G. A., St. George-Hyslop, P., Van Keuren, M. L., Patterson, D., Pagan, S., Kurnit, D. M. & Neve, R. L. (1987) *Science* **235**, 880–884.
5. Goldgaber, D., Lerman, M. I., McBride, O. W., Saffiotti, U. & Gajdusek, D. C. (1987) *Science* **235**, 877–880.
6. Robakis, N. K., Ramakrishna, N., Wolfe, G. & Wisniewski, H. M. (1987) *Proc. Natl. Acad. Sci. USA* **84**, 4190–4194.
7. Levy, E., Carman, M., Fernandez-Madrid, I., Power, M., Lieberburg, I., van Duinen, S., Gerard, T., Bots, A., Luyendijk, W. & Frangione, B. (1990) *Science* **248**, 1124–1128.
8. Goate, A., Chartie-Harlin, M. C., Mullan, M., Brown, J., Crawford, F., *et al.* (1991) *Nature (London)* **349**, 704–706.
9. Murrell, J., Farlow, M., Ghetti, B. & Benson, M. D. (1991) *Science* **254**, 97–99.
10. Mullan, M., Crawford, F., Axelman, K., Houlden, H., Lilius, L., Winblad, B. & Lannfelt, L. (1992) *Nat. Genet.* **1**, 345–347.
11. Jones, C. T., Morris, S., Yates, C. M., Moffoot, A., Sharpe, C., Brock, D. J. & St. Clair, D. (1992) *Nat. Genet.* **1**, 306–309.
12. Hendriks, L., van Duijn, C. M., Cras, P., Cruts, M., Van Hul, W., van Harskamp, F., Warren, A., McInnis, M. G., Antonarakis, S. E., Martin, J. J., Hofman, A. & Van Broeckhaven, C. (1992) *Nat. Genet.* **1**, 218–221.
13. Citron, M., Oltersdorf, T., Haass, C., McConlogue, L., Hung, A. Y., Seubert, P., Vigo-Pelfrey, C., Lieberburg, I. & Selkoe, D. J. (1992) *Nature (London)* **360**, 672–674.
14. Cai, X. D., Golde, T. E. & Younkin, S. G. (1993) *Science* **259**, 514–516.
15. Shoji, M., Golde, T. E., Ghiso, J., Cheung, T. T., Estus, S., Shaffer, L. M., Cai, X.-D., McKay, D. M., Tintner, R., Frangione, B. & Younkin, S. G. (1992) *Science* **258**, 126–129.
16. Seubert, P., Vigo-Pelfrey, C., Esch, F., Lee, M., Dovey, H., Davis, D., Sinha, S., Schlossmacher, M., Whaley, J., Swindlehurst, C., McCormack, R., Wolfert, R., Selkoe, D., Lieberburg, I. & Shenk, D. (1992) *Nature (London)* **359**, 325–327.
17. Haass, C., Schlossmacher, M. G., Hung, A. Y., Vigo-Pelfrey, C., Mellon, A., Ostaszewski, B. L., Lieberburg, I., Koo, E. H., Schenk, D., Teplow, D. B. & Selkoe, D. J. (1992) *Nature (London)* **359**, 322–325.
18. van Gool, W. A., Schenk, D. B. & Bolhuis, P. A. (1994) *Neurosci. Lett.* **172**, 122–124.
19. Miller, D. L., Papayannopoulos, I. A., Styles, J., Bobin, S. A., Lin, Y. Y., Biemann, K. & Iqbal, K. (1993) *Arch. Biochem. Biophys.* **301**, 41–52.
20. Mori, H., Takio, K., Ogawara, M. & Selkoe, D. J. (1992) *J. Biol. Chem.* **267**, 17082–17086.
21. Iwatsubo, T., Odaka, A., Suzuki, N., Mizusawa, H., Nukina, N. & Ihara, Y. (1994) *Neuron* **13**, 45–53.
22. Jarrett, J. T., Berger, E. P. & Lansbury, P. T., Jr. (1993) *Biochemistry* **32**, 4693–4697.
23. Jarrett, J. T. & Lansbury, P. T., Jr. (1993) *Cell* **73**, 1055–1058.
24. Kelly, J. W. & Lansbury, P. T., Jr. (1994) *Amyloid Int. J. Exp. Clin. Invest.* **1**, 186–205.
25. Pike, C. J., Walencewicz-Wasserman, J., Kosmoski, J., Cribbs, D. H., Glabe, C. G. & Cotman, C. W. (1995) *J. Neurochem.* **64**, 253–265.
26. Ma, J., Yee, A., Brewer, H., Jr., Das, S. & Potter, H. (1994) *Nature (London)* **372**, 92–94.
27. Ausubel, F. M., Brent, R., Kingston, R. E., Moore, D. D., Seidman, J. G., Smith, J. A. & Struhl, K., eds. (1994) *Current Protocols in Molecular Biology* (Wiley, New York), pp. 13.6.2–13.7.5.
28. Sambrook, J., Fritsch, E. F. & Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Lab. Press, Plainview, NY).
29. Golemis, E. A. & Brent, R. (1992) *Mol. Cell Biol.* **12**, 3006–3014.
30. Ma, J. & Ptashne, M. (1987) *Cell* **51**, 113–119.
31. Gill, G. & Ptashne, M. (1988) *Nature (London)* **334**, 721–724.
32. Harlow, E. & Lane, D. (1988) *Antibodies: A Laboratory Manual* (Cold Spring Harbor Lab. Press, Plainview, NY), 2nd Ed.
33. Brent, R. & Ptashne, M. (1984) *Nature (London)* **312**, 612–615.
34. Samson, J.-L., Jacson-Grusby, L. & Brent, R. (1989) *Cell* **57**, 1045–1052.
35. Gyuris, J., Golemis, E., Chertkov, H. & Brent, R. (1993) *Cell* **75**, 791–803.
36. Hilbich, C., Kisters Woike, B., Reed, J., Masters, C. L. & Beyreuther, K. (1992) *J. Mol. Biol.* **228**, 460–473.
37. Kocisco, D. A., Corne, J. H., Priola, S. A., Cheseboro, B., Raymond, G. J., Lansbury, P. T. & Caughey, B. (1994) *Nature (London)* **370**, 471–474.
38. Hawkins, P. N., Cavender, J. P & Pepys, M. B. (1990) *N. Engl. J. Med.* **323**, 508–513.
39. Colon, W. & Kelly, J. W. (1992) *Biochemistry* **31**, 8654–8660.
40. Jarrett, J. T. & Lansbury, P. T., Jr. (1992) *Biochemistry* **31**, 12345–12352.
41. Come, J. H., Fraser, P. E. & Lansbury, P. T., Jr. (1993) *Proc. Natl. Acad. Sci. USA* **90**, 5959–5963.

US005854204A

United States Patent [19]

Findeis et al.

[11] Patent Number: 5,854,204

[45] Date of Patent: Dec. 29, 1998

[54] Aβ PEPTIDES THAT MODULATE β-AMYLOID AGGREGATION

[75] Inventors: **Mark A. Findeis**, Cambridge; **Howard Benjamin**, Lexington; **Marc B. Garnick**, Brookline; **Malcolm L. Gefter**, Lincoln; **Arvind Hundal**, Brighton, all of Mass.; **Laura Kasman**, Athens, Ga.; **Gary Musso**, Hopkinton, Mass.; **Ethan R. Signer**, Cambridge, Mass.; **James Wakefield**, Brookline, Mass.; **Michael Reed**, Marietta, Ga.; **Susan Molineaux**, Brookline, Mass.; **William Kubasek**, Belmont, Mass.; **Joseph Chin**, Salem, Mass.; **Jung-Ja Lee**, Wayland, Mass.; **Michael Kelley**, Arlington, Mass.

[73] Assignee: **Praecis Pharmaceuticals, Inc.**, Cambridge, Mass.

[21] Appl. No.: **612,785**

[22] Filed: **Mar. 14, 1996**

Related U.S. Application Data

[63] Continuation-in-part of Ser. No. 404,831, Mar. 14, 1995, and a continuation-in-part of Ser. No. 475,579, Jun. 7, 1995, and a continuation-in-part of Ser. No. 548,998, Oct. 27, 1995.

[51] **Int. Cl.**[6] **C07K 14/435**; C07K 7/08

[52] **U.S. Cl.** **514/2**; 514/12; 514/14; 530/324; 530/326

[58] **Field of Search** 514/14, 12, 2; 530/300, 324, 326; 930/10

[56] **References Cited**

U.S. PATENT DOCUMENTS

5,338,663	8/1994	Potter et al.	435/4
5,470,951	11/1995	Roberts	530/330

FOREIGN PATENT DOCUMENTS

554 887 A1	8/1993	European Pat. Off. .
641 861 A1	3/1995	European Pat. Off. .
681 844 A1	11/1995	European Pat. Off. .
WO 93/04194	3/1993	WIPO .
WO 94/28412	12/1994	WIPO .
WO 95/05394	2/1995	WIPO .
WO 95/05604	2/1995	WIPO .
WO 95/12815	5/1995	WIPO .
WO 95/20979	8/1995	WIPO .

OTHER PUBLICATIONS

Barrow, Colin J. and Michael G. Zagorski (1991) "Solution Structures of β Peptide and Its Constituent Fragments: Relation to Amyloid Deposition" *Science* 253: 179–182.

Barrow, Colin J. et al. (1992) "Solution Conformations and Aggregational Properties of Synthetic Amyloid β–Peptides of Alzheimer's Disease: Analysis of Circular Dichroism Spectra" *J. Mol. Biol.* 225: 1075–1093.

Brown, Abraham M. et al. (1994) "Biotinylated and Cysteine–Modified Peptides as Useful Reagents for Studying the Inhibition of Cathepsin G" *Analytical Biochemistry* 217: 139–147.

Burdick, Debra et al. (1992) "Assembly and Aggregation Properties of Synthetic Alzheimer's A4/β Amyloid Peptide Analogs" *Journal of Biological Chemistry* 267(1): 546–554.

Chantry, Andrew et al. (1992) "Biotinyl Analogues of Amylin as Biologically Active Probes for Amylin/CGRP Receptor Recognition" *FEBS* 296(2): 123–127.

Clements, Angela et al. (1993) "Aggregation of Alzheimer's Peptides" *Biochemical Society Transactions* 22: 16S.

Come, Jon H. et al. (1993) "A Kinetic Model for Amyloid Formation in the Prion Diseases: Importance of Seeding" *Proc. Natl. Acad. Sci. USA* 90: 5959–5963.

Evans, Krista C. et al. (1995) "Apolipoprotein E Is a Kinetic But Not a Thermodynamic Inhibitor of Amyloid Formation: Implications for the Pathogenesis and Treatment of Alzheimer Disease" *Proc. Natl. Acad. Sci. USA* 92: 763–767.

Fabian, Heinz et al. (1993) "Comparative Analysis of Human and Dutch–Type Alzheimer β–Amyloid Peptides by Infrared Spectroscopy and Circular Dichroism" *Biochemical and Biophysical Research Communications* 191(1): 232–239.

Fabian, Heinz et al. (1994) "Synthetic Post–Translationally Modified Human Aβ Peptide Exhibits a Markedly Increased Tendency to Form β–Pleated Sheets in vitro" *Eur. J. Biochem.* 221: 959–964.

Flood, J.F. et al., (1994) "Topography of a Binding Site for Small Amnestic Peptides Deduced from Structure–Activity Studies: Relation to Amnestic Effect of Amyloid β Protein," *Proc. Natl. Acad. Sci. USA* vol. 91, pp. 380–384.

Fraser, Paul E. et al. (1994) "Conformation and Fibrillogenesis of Alzheimer Aβ Peptides with Selected Substitution of Charged Residues" *J. Mol. Biol.* 244: 64–73.

Fraser, Paul E. et al. (1992) "Fibril Formation by Primate, Rodent, and Dutch–Hemorrhagic Analogues of Alzheimer Amyloid β–Protein" *Biochemistry* 31: 10716–10723.

(List continued on next page.)

Primary Examiner—Paula K. Hutzell
Assistant Examiner—Robert C. Hayes
Attorney, Agent, or Firm—Lahive & Cockfield, LLP; Guilio A. DeConti, Jr.; Catherine J. Kara

[57] **ABSTRACT**

Compounds that modulate the aggregation of amyloidogenic proteins or peptides are disclosed. The modulators of the invention can promote amyloid aggregation or, more preferably, can inhibit natural amyloid aggregation. In a preferred embodiment, the compounds modulate the aggregation of natural β amyloid peptides (β-AP). In a preferred embodiment, the β amyloid modulator compounds of the invention are comprised of an Aβ aggregation core domain and a modifying group coupled thereto such that the compound alters the aggregation or inhibits the neurotoxicity of natural β amyloid peptides when contacted with the peptides. Furthermore, the modulators are capable of altering natural β-AP aggregation when the natural β-APs are in a molar excess amount relative to the modulators. Pharmaceutical compositions comprising the compounds of the invention, and diagnostic and treatment methods for amyloidogenic diseases using the compounds of the invention, are also disclosed.

10 Claims, 7 Drawing Sheets

OTHER PUBLICATIONS

Gorevic, PD et al. (1987) "Ten to Fourteen Residue Peptides of Alzheimer's Disease Protein are Sufficient for Amyloid Fibril Formation and Its Characteristic Xray Diffraction Pattern" *Biochemical and Biophysical Research Communications* 147(2): 854–862.

Gowing, Eric et al. (1994) "Chemical Characterization of Aβ 17–42 Peptide, a Component of Diffuse Amyloid Deposits of Alzheimer Disease" *J. Biol. Chem.* 269(15): 10987–10990.

Halverson, Kurt et al. (1990) "Molecular Determinants of Amyloid Deposition in Alzheimer's Disease: Conformational Studies of Synthetic β–Protein Fragments" *Biochemistry* 29(11): 2639–2644.

Hansen, Morten B. et al. (1989) "Re–examination and Further Development of a Precise and Rapid Dye Method for Measuring Cell Growth/Cell Kill" *J. Immunol. Meth.* 119: 203–210.

Hardy, John A. and Gerald A. Higgins (1992) "Alzheimer's Disease: The Amyloid Cascade Hypothesis" *Science* 256: 184–185.

Hilbich, Caroline et al. (1991) "Aggregation and Secondary Structure of Synthetic Amyloid βA4 Paptides of Alzheimer's Disease" *J. Mol. Biol.* 218: 149–163.

Hilbich, Caroline et al. (1991) "Human and Rodent Sequence Analogs of Alzheimer's Amyloid βA4 A4 Share Similar Properties and can Be Solubilized in Buffers of pH 7.4" *Eur. J. Biochem.* 201: 61–69.

Hilibich, Caroline et al. (1992) "Substitutions of Hydrophobic Amino Acids Reduce the Amyloidogenicity of Alzheimer's Disease βA4 Peptides" *J. Mol. Biol.* 228: 460–473.

Inouye, H. et al. (1993) "Structure of Beta–Crystallite Assemblies Formed by Alzheimer β–Amyloid Protein Analogs: Analysis by X–ray Diffraction," *Chemical Abstracts* vol. 119, p. 349, Abstract No. 119.

Jarrett, Joseph T. and Peter T. Lansbury, Jr. (1993) "Seeding 'One–Dimensional Crystallization' of Amyloid: A Pathogenic Mechanism in Alzheimer's Disease and Scrapie?" *Cell* 73: 1055–1058.

Jarrett, Joseph T. et al. (1993) "The Carboxy Terminus of the β Amyloid Protein Is Critical for the Seeding of Amyloid Formation: Implications for the Pathogenesis of Alzheimer's Disease" *Biochemistry* 32(18): 4693–4697.

Jarrett, Joseph T. et al. (1994) "Models of the β Protein C–Terminus: Differences in Amyloid Structure May Lead to Segregation of 'Long' and 'Short' Fibrils" *Journal of the American chemical Society* 116(21): 9741–9742.

Kelly, Jeffrey W. and Peter T. Lansbury, Jr. (1994) "A Chemical Approach to Elucidate the Mechanism of Transthyretin and β–Protein Amyloid Fibril Formation" *Int. J. Exp. Clin. Invest* 1: 186–205.

Kirschner, Daniel A. et al. (1987) "Synthetic Peptide Homologous to β Protein from Alzheimer Disease forms Amyloid–like Fibrils in vitro" *Proc. Natl. Acad. Sci. USA* 84: 6953–6957.

Klunk, William E. and Jay W. Pettegrew (1990) "Alzheimer's β–Amyloid Protein Is Covalently Modified When Dissolved in Formic Acid" *Journal of Neurochemistry* 54(6): 2050–2054.

Lansbury, Jr., Peter T. (1992) "In Pursuit of the Molecular Structure of Amyloid Plaque: New Technology Provides Unexpected and Critical Information" *Biochemistry* 31(30): 6866–6870.

LeVine, III, Harry (1993) "Thioflavine T Interaction with Synthetic Alzheimer's Disease β–Amyloid Peptides: Detection of Amyloid Aggregtion in Solution" *Protein Science* 2: 404–410.

Maggio, John E. et al. (1992) "Reversible in vitro Growth of Alzheimer Disease β–Amyloid Plaques by Deposition of Labeled Amyloid Peptide" *Proc. Natl. Acad. Sci. USA* 89: 5462–5466.

Miller, Brian T. et al. (1994) "Identification and Characterization of O–Biotinylated Hydroxy Amino Acid Residues in Peptides" *Analytical Biochemistry* 219: 240–248.

Orlando, Ron et al. (1992) "Covalent Modification of Alzheimer's Amyloid β– Peptide in Formic Acid Solutions" *Biochemical and Biophysical Research Communications* 184(2): 686–691.

Pike, Christian J. et al. (1993) "Neurodegeneration induced by β–Amyloid Peptides in vitro: The Role of Peptide Assembly State" *Journal of Neuroscience* 13(4): 1676–1687.

Pike, Christian J. et al. (1995) "Structure–Activity Analyses of β–Amyloid Peptides: Contributions of the β25–35 Region to Aggregation and Neurotoxity" *Journal of Neurochemistry* 64(1): 253–265.

Schwarzman, Alexander L. et al. (1994) "Transthyretin Sequesters Amyloid β Protein and Prevents Amyloid Formation" *Proc. Natl. Acad. Sci. USA* 91: 8368–8372.

Shearman, Mark S. et al. (1994) "Inhibition of PC12 Cell Redox Activity is a Specific, Early Indicator of the Mechanism of β–Amyloid–Mediated Cell Death" *Proc. Natl. Acad. Sci. USA* 91: 1470–1474.

Shen, Chih–Lung et al. (1994) "Effect of Acid Predissolution on Fibril Size and Fibril Flexibility of Synthetic β–Amyloid Peptide" *Biophysical Journal* 67: 1238–1246.

Shen, Chih–Lung et al. (1993) "Light Scattering Analysis of Fibril Growth from the Amino–Terminal Fragment β(1–28) of β–Amyloid Peptide" *Biophysical Journal* 65: 2383–2395.

Snyder, Seth W. et al. (1994) "Amyloid–β Aggregation: Selective Inhibition of Aggregation in Mixtures of Amyloid with Different Chain Lengths" *Biophysical Journal* 67: 1216–1228.

Sonnenberg–Reines, J. et al. (1993) "Biotinylated and Cysteine Modified Peptides as Useful Reagents for Studying the Inhibition of Putative N–terminal B–Amyloid Peptide Enzymes," *Society for Neuroscience Abstracts* vol. 19 (1–3), p. 861.

Soreghan, Brian et al. (1994) "Surfactant Properties of Alzheimer's Aβ Peptides and the Mechanism of Amyloid Aggregation" *The Journal of Biological Chemistry* 158(46): 28551–28554.

Sorimachi, Kay and David J. Craik (1994) "Structure Determination of Extracellular Fragments of Amyloid Proteins Involved in Alzheimer's Disease and Dutch–type Hereditary Cerebral Haemorrhage with Amyloidosis" *Eur. J. Biochem* 219: 237–251.

Strittmatter, Warren J. et al. (1993) "Binding of Human Apolipoprotein E to Synthetic Amyloid β Peptide: Isoform–Specific Effects and Implications for Late–Onset Alzheimer Disease" *Proc. Natl. Acad. Sci. USA* 90: 8098–8102.

Tomiyama, Takami et al. (1994) "Racemization of Asp^{23} Residue Affects the Aggregation Properties of Alzheimer Amyloid β Protein Analogues" *J. Biol. Chem.* 269(14): 10205–10208.

Tomski, Sharon J. and Regina M. Murphy (1992) "Kinetics of Aggregation of Synthetic β–Amyloid Peptide" *Archives of Biochemistry and Biophysics* 294(2): 630–638.

Vitek, Michael P. et al. (1994) "Advanced Glycation End Products Contribute to Amyloidosis in Alzheimer Disease" *Proc. Natl. Acad. Sci. USA* 91: 4766–4770.

Vyas, S. B. et al. "Characterization of Aggregation in Alzheimer β–protein Using Synthetic Peptide Fragments on Reverse–Phase Matrix," in *Peptides, Chemistry and Biology* (J.A. Smith and J.E. Rivier, eds.), ESCOM, Leiden, 1992, pp. 278–279.

Weinreb, Paul H. et al. (1994) "Peptide Models of a Hydrophobic Cluster at the C–Terminus of the β–Amyloid" *Journal of the American Chemical Society* 116(23): 10835–10836.

Woods, S. J. et al. (1995) "Prolines and Amyloidogenicity in Fragments of the Alzheimer's Peptide β/A4," *Biochemistry* vol. 34, 724–730.

Rudinger, In "Peptide Hormones", Jun. 1976, ed. J.A. Parsons, University Park Press, Baltimore, pp. 1–7.

FIG.1

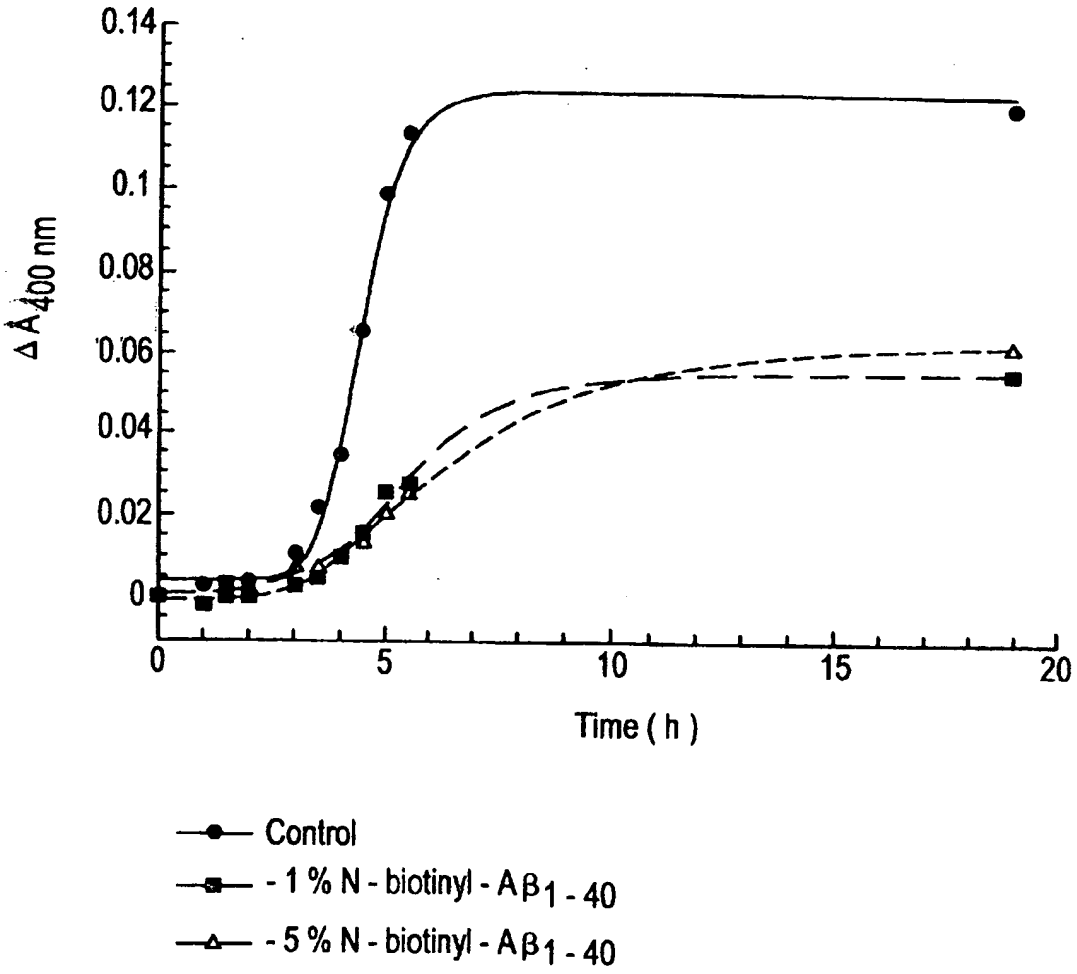
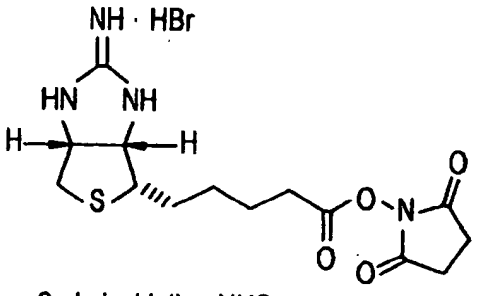
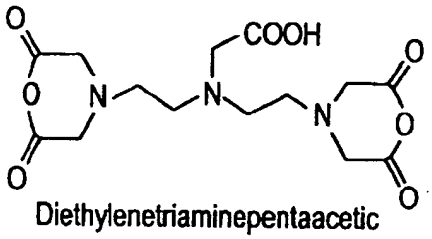

0.14
0.12
0.1
0.08
0.06
0.04
0.02
0
ΔA400 nm
0
5
10
15
20
Time (h)
Control
- 1 % N - biotinyl - Aβ1 - 40
- 5 % N - biotinyl - Aβ1 - 40

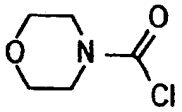
FIG.2A



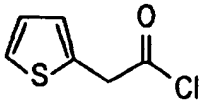
2 - Iminobiotin - NHS ester



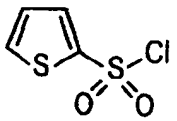
Diethylenetriaminepentaacetic Dianhydride



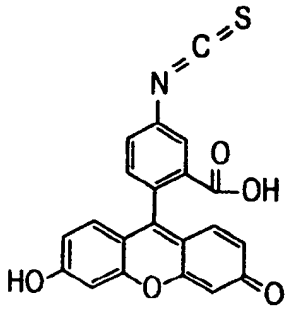
4 - Morpholinecarbonyl Chloride



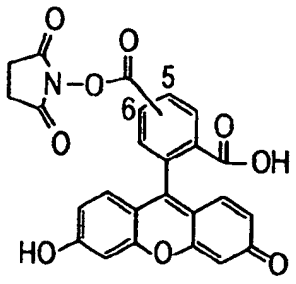
2 - Thiopheneacetyl Chloride



2 - Thiophenesulfonyl Chloride

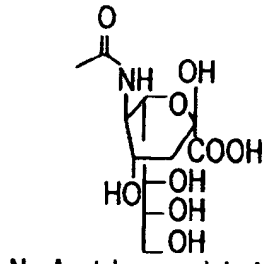


Fluorescein Isothiocyanate Isomer I

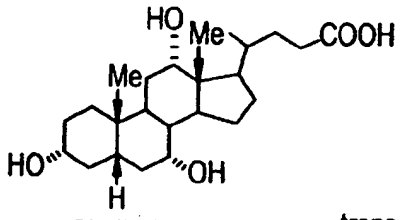


5 - (and - 6) - Carboxyfluorescein, succinimidyl ester

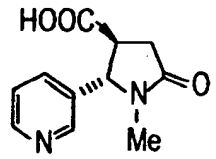
FIG.2B



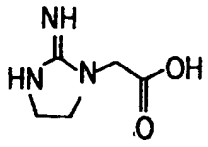
N - Acetylneuraminic Acid



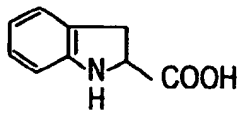
Cholic Acid



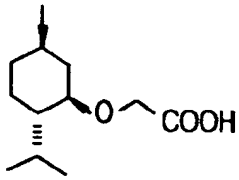
trans - 4 - Cotininecarboxylic Acid



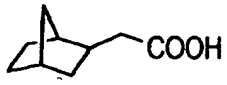
2 - Imino - 1 - Imidazolidineacetic Acid



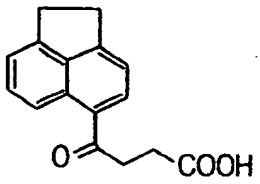
(S) - (-) - Indoline - 2 - carboxylic Acid



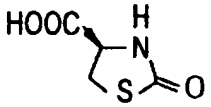
(-) - Menthoxyacetic Acid



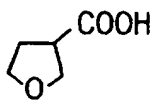
2 - Norbornaneacetic Acid



γ - Oxo - 5 - acenaphthenebutyric Acid



(-) - 2 - Oxo - 4 - thiazolidine - carboxylic Acid



Tetrahydro - 3 - furoic Acid

FIG.3

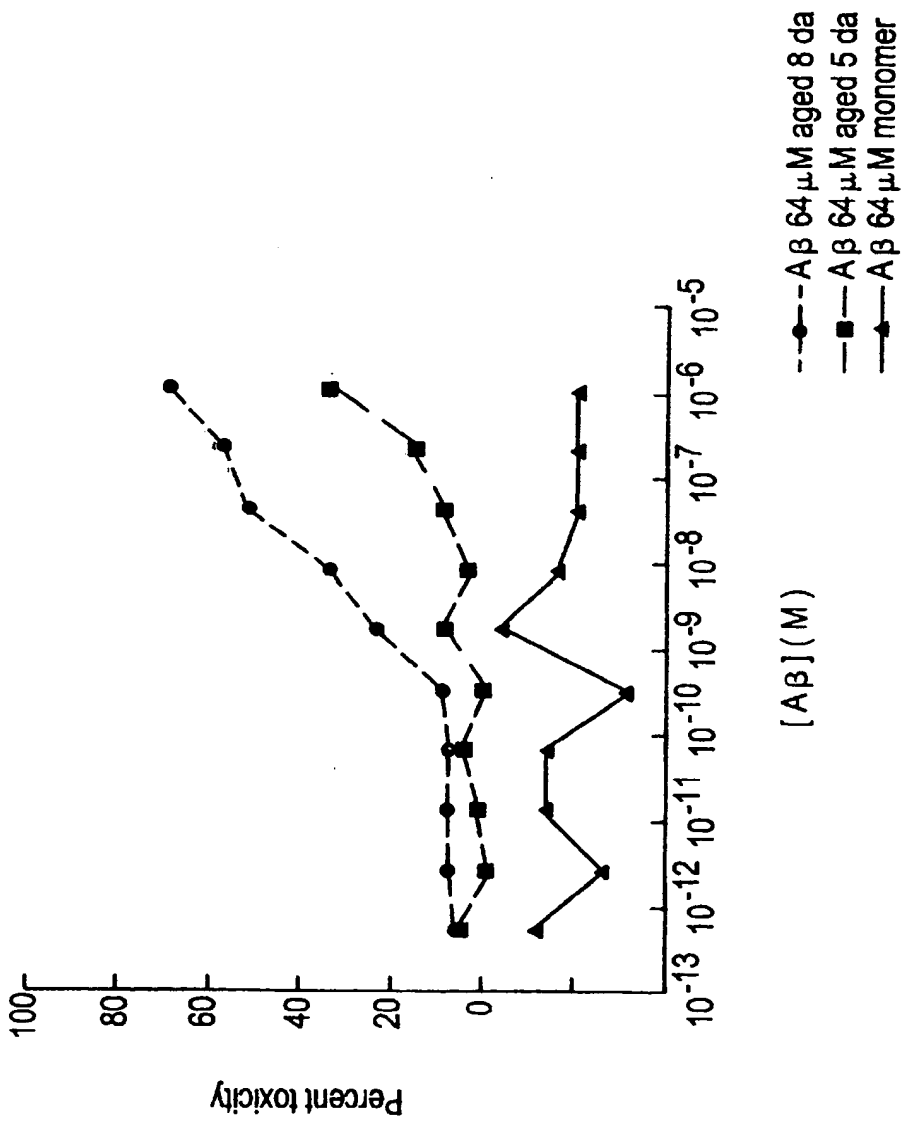

Percent toxicity
100
80
60
40
20
0
10^{-13} 10^{-12} 10^{-11} 10^{-10} 10^{-9} 10^{-8} 10^{-7} 10^{-6} 10^{-5}
[Aβ](M)
Aβ 64 μM aged 8 da
Aβ 64 μM aged 5 da
Aβ 64 μM monomer

FIG.4B

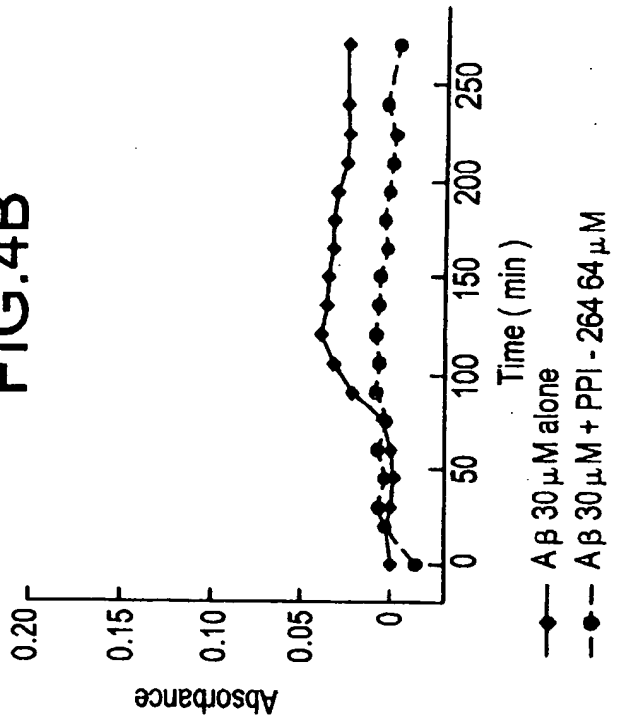

Absorbance
0.20
0.15
0.10
0.05
0
0
50
100
150
200
250
Time (min)
Aβ 30 μM alone
Aβ 30 μM + PPI - 264 64 μM

FIG.4A

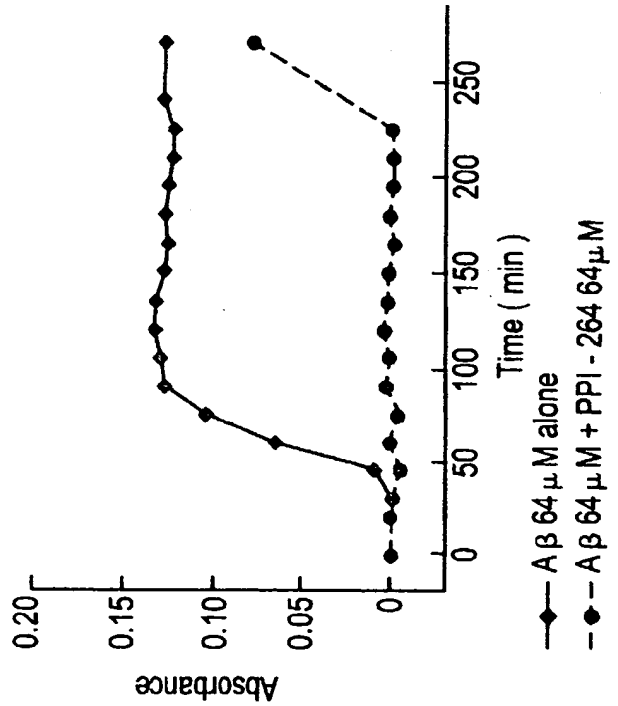

Absorbance
0.20
0.15
0.10
0.05
0
0
50
100
150
200
250
Time (min)
Aβ 64 μM alone
Aβ 64 μM + PPI - 264 64μM

FIG.4D

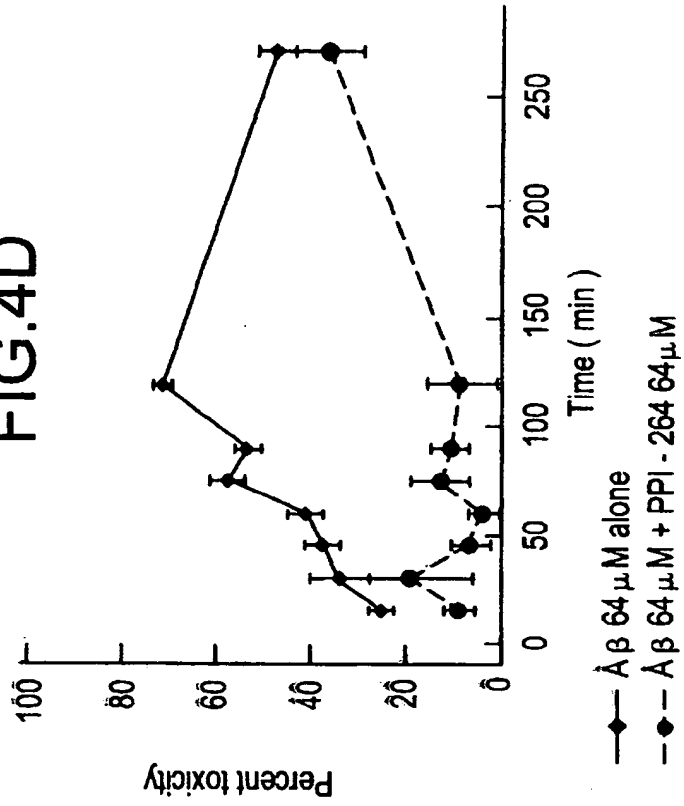

Percent toxicity
100
80
60
40
20
0
0
50
100
150
200
250
Time (min)
Aβ 64 μM alone
Aβ 64 μM + PPI - 264 64μM

FIG.4C

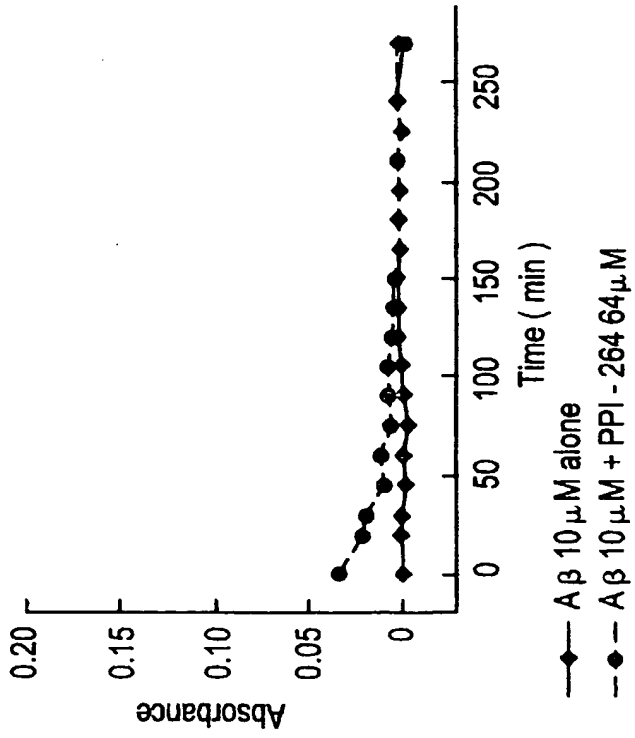

Absorbance
0.20
0.15
0.10
0.05
0
0
50
100
150
200
250
Time (min)
Aβ 10 μM alone
Aβ 10 μM + PPI - 264 64μM

FIG.4F

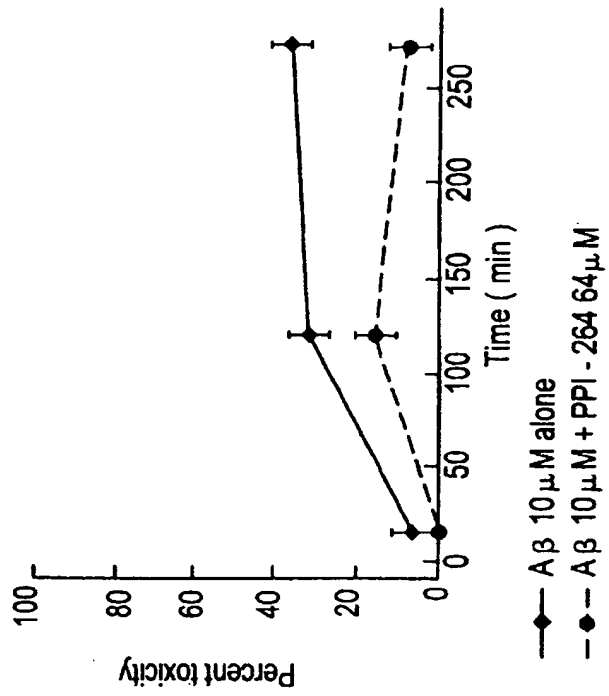
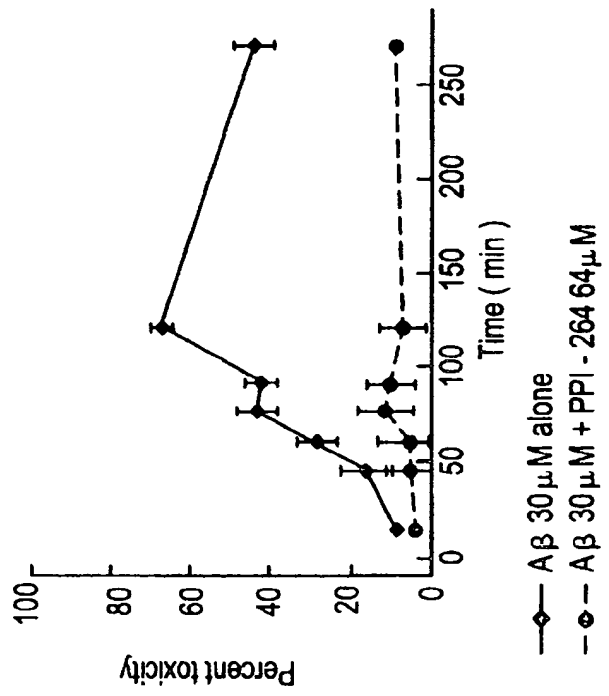

Percent toxicity
0
20
40
60
80
100
0
50
100
150
200
250
Time (min)
Aβ 10 μM alone
Aβ 10 μM + PPI - 264 64μM

FIG.4E


Percent toxicity
0
20
40
60
80
100
0
50
100
150
200
250
Time (min)
Aβ 30 μM alone
Aβ 30 μM + PPI - 264 64μM

Aβ PEPTIDES THAT MODULATE β-AMYLOID AGGREGATION

RELATED APPLICATIONS

This application is a continuation-in-part of U.S. patent application Ser. No. 08/404,831, filed Mar. 14, 1995, and U.S. patent application Ser. No. 08/475,579, filed Jun. 7, 1995 and U.S. patent application Ser. No. 08/548,998, filed Oct. 27, 1995, the entire contents of each of which are hereby incorporated by reference.

BACKGROUND OF THE INVENTION

Alzheimer's disease (AD), first described by the Bavarian psychiatrist Alois Alzheimer in 1907, is a progressive neurological disorder that begins with short term memory loss and proceeds to disorientation, impairment of judgment and reasoning and, ultimately, dementia. The course of the disease usually leads to death in a severely debilitated, immobile state between four and 12 years after onset. AD has been estimated to afflict 5 to 11 percent of the population over age 65 and as much as 47 percent of the population over age 85. The societal cost for managing AD is upwards of 80 billion dollars annually, primarily due to the extensive custodial care required for AD patients. Moreover, as adults born during the population boom of the 1940's and 1950's approach the age when AD becomes more prevalent, the control and treatment of AD will become an even more significant health care problem. Currently, there is no treatment that significantly retards the progression of the disease. For reviews on AD, see Selkoe, D. J. *Sci. Amer.*, November 1991, pp. 68–78; and Yankner, B. A. et al. (1991) *N. Eng. J. Med.* 325:1849–1857.

It has recently been reported (Games et al. (1995) *Nature* 373:523–527) that an Alzheimer-type neuropathology has been created in transgenic mice. The transgenic mice express high levels of human mutant amyloid precursor protein and progressively develop many of the pathological conditions associated with AD.

Pathologically, AD is characterized by the presence of distinctive lesions in the victim's brain. These brain lesions include abnormal intracellular filaments called neurofibrillary tangles (NTFs) and extracellular deposits of amyloidogenic proteins in senile, or amyloid, plaques. Amyloid deposits are also present in the walls of cerebral blood vessels of AD patients. The major protein constituent of amyloid plaques has been identified as a 4 kilodalton peptide called β-amyloid peptide (β-AP)(Glenner, G. G. and Wong, C. W. (1984) *Biochem. Biophys. Res. Commun.* 120:885–890; Masters, C. et al. (1985) *Proc. Natl. Acad. Sci. USA* 82:4245–4249). Diffuse deposits of β-AP are frequently observed in normal adult brains, whereas AD brain tissue is characterized by more compacted, dense-core β-amyloid plaques. (See e.g., Davies, L. et al. (1988) *Neurology* 38:1688–1693). These observations suggest that β-AP deposition precedes, and contributes to, the destruction of neurons that occurs in AD. In further support of a direct pathogenic role for β-AP, β-amyloid has been shown to be toxic to mature neurons, both in culture and in vivo. Yankner, B. A. et al. (1989) *Science* 245:417–420; Yankner, B. A. et al. (1990) *Proc. Natl. Acad. Sci. USA* 87:9020–9023; Roher, A. E. et al. (1991) *Biochem. Biophys. Res. Commun.* 174:572–579; Kowall, N. W. et al. (1991) *Proc. Natl. Acad. Sci. USA* 88:7247–7251. Furthermore, patients with hereditary cerebral hemorrhage with amyloidosis-Dutch-type (HCHWA-D), which is characterized by diffuse β-amyloid deposits within the cerebral cortex and cerebrovasculature,

have been shown to have a point mutation that leads to an amino acid substitution within β-AP. Levy, E. et al. (1990) *Science* 248:1124–1126. This observation demonstrates that a specific alteration of the β-AP sequence can cause β-amyloid to be deposited.

Natural β-AP is derived by proteolysis from a much larger protein called the amyloid precursor protein (APP). Kang, J. et al. (1987) *Nature* 325:733; Goldgaber, D. et al. (1987) *Science* 235:877; Robakis, N. K. et al. (1987) *Proc. Natl. Acad. Sci. USA* 84:4190; Tanzi, R. E. et al. (1987) *Science* 235:880. The APP gene maps to chromosome 21, thereby providing an explanation for the β-amyloid deposition seen at an early age in individuals with Down's syndrome, which is caused by trisomy of chromosome 21. Mann, D. M. et al. (1989) *Neuropathol. Appl. Neurobiol.* 15:317; Rumble, B. et al. (1989) *N. Eng. J Med.* 320:1446. APP contains a single membrane spanning domain, with a long amino terminal region (about two-thirds of the protein) extending into the extracellular environment and a shorter carboxy-terminal region projecting into the cytoplasm. Differential splicing of the APP messenger RNA leads to at least five forms of APP, composed of either 563 amino acids (APP-563), 695 amino acids (APP-695), 714 amino acids (APP-714), 751 amino acids (APP-751) or 770 amino acids (APP-770).

Within APP, naturally-occurring β amyloid peptide begins at an aspartic acid residue at amino acid position 672 of APP-770. Naturally-occurring β-AP derived from proteolysis of APP is 39 to 43 amino acid residues in length, depending on the carboxy-terminal end point, which exhibits heterogeneity. The predominant circulating form of β-AP in the blood and cerebrospinal fluid of both AD patients and normal adults is β1–40 ("short β"). Seubert, P. et al. (1992) *Nature* 359:325; Shoji, M. et al. (1992) *Science* 258:126. However, β1–42 and β1–43 ("long β") also are forms in β-amyloid plaques. Masters, C. et al. (1985) *Proc. Natl Acad. Sci. USA* 82:4245; Miller, D. et al. (1993) *Arch. Biochem. Biophys.* 301:41; Mori, H. et al. (1992) *J. Biol. Chem.* 267:17082. Although the precise molecular mechanism leading to β-APP aggregation and deposition is unknown, the process has been likened to that of nucleation-dependent polymerizations, such as protein crystallization, microtubule formation and actin polymerization. See e.g., Jarrett, J. T. and Lansbury, P. T. (1993) *Cell* 73:1055–1058. In such processes, polymerization of monomer components does not occur until nucleus formation. Thus, these processes are characterized by a lag time before aggregation occurs, followed by rapid polymerization after nucleation. Nucleation can be accelerated by the addition of a "seed" or preformed nucleus, which results in rapid polymerization. The long β forms of β-AP have been shown to act as seeds, thereby accelerating polymerization of both long and short β-AP forms. Jarrett, J. T. et al. (1993) *Biochemistry* 32:4693.

In one study, in which amino acid substitutions were made in β-AP, two mutant β peptides were reported to interfere with polymerization of non-mutated β-AP when the mutant and non-mutant forms of peptide were mixed. Hilbich, C. et al. (1992) *J. Mol. Biol.* 228:460–473. However, equimolar amounts of the mutant and non-mutant (i.e., natural) β amyloid peptides were used to see this effect and the mutant peptides were reported to be unsuitable for use in vivo. Hilbich, C. et al. (1992), supra.

SUMMARY OF THE INVENTION

This invention pertains to compounds, and pharmaceutical compositions thereof, that can modulate the aggregation of amyloidogenic proteins and peptides, in particular com-

pounds that can modulate the aggregation of natural β amyloid peptides (β-AP) and inhibit the neurotoxicity of natural β-APs. In one embodiment, the invention provides an amyloid modulator compound comprising an amyloidogenic protein, or peptide fragment thereof, coupled directly or indirectly to at least one modifying group such that the compound modulates the aggregation of natural amyloid proteins or peptides when contacted with the natural amyloidogenic proteins or peptides. Preferably, the compound inhibits aggregation of natural amyloidogenic proteins or peptides when contacted with the natural amyloidogenic proteins or peptides. The amyloidogenic protein, or peptide fragment thereof, can be, for example, selected from the group consisting of transthyretin (TTR), prion protein (PrP), islet amyloid polypeptide (IAPP), atrial natriuretic factor (ANF), kappa light chain, lambda light chain, amyloid A, procalcitonin, cystatin C, β2 microglobulin, ApoA-I, gelsolin, calcitonin, fibrinogen and lysozyme.

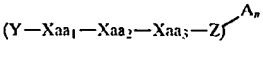
In the most preferred embodiment of the invention, the compound modulates the aggregation of natural β-AP. The invention provides a β-amyloid peptide compound comprising a formula:

wherein Xaa is a β-amyloid peptide having an amino-terminal amino acid residue corresponding to position 668 of β-amyloid precursor protein-770 (APP-770) or to a residue carboxy-terminal to position 668 of APP-770, A is a modifying group attached directly or indirectly to the β-amyloid peptide of the compound such that the compound inhibits aggregation of natural β-amyloid peptides when contacted with the natural β-amyloid peptides, and n is an integer selected such that the compound inhibits aggregation of natural β-amyloid peptides when contacted with the natural β-amyloid peptides.

In one embodiment, at least one A group is attached directly or indirectly to the amino terminus of the β-amyloid peptide of the compound. In another embodiment, at least one A group is attached directly or indirectly to the carboxy terminus of the β-amyloid peptide of the compound. In yet another embodiment, at least one A group is attached directly or indirectly to a side chain of at least one amino acid residue of the β-amyloid peptide of the compound.

The invention also provides a β-amyloid modulator compound comprising an Aβ aggregation core domain (ACD) coupled directly or indirectly to at least one modifying group (MG) such that the compound modulates the aggregation or inhibits the neurotoxicity of natural β-amyloid peptides when contacted with the natural β-amyloid peptides. Preferably, the Aβ aggregation core domain is modeled after a subregion of natural β-amyloid peptide between 3 and 10 amino acids in length.

The invention also provides β-amyloid modulator compound comprising a formula:



wherein Xaa_1, Xaa_2 and Xaa_3 are each amino acid structures and at least two of Xaa_1, Xaa_2 and Xaa_3 are, independently, selected from the group consisting of a leucine structure, a phenylalanine structure and a valine structure;

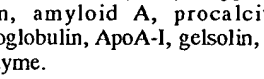
Y, which may or may not be present, is a peptidic structure having the formula $(Xaa)_a$, wherein Xaa is any amino acid structure and a is an integer from 1 to 15;

Z, which may or may not be present, is a peptidic structure having the formula $(Xaa)_b$, wherein Xaa is any amino acid structure and b is an integer from 1 to 15; and

A is a modifying group attached directly or indirectly to the compound and n is an integer;

Xaa_1, Xaa_2, Xaa_3, Y, Z, A and n being selected such that the compound modulates the aggregation or inhibits the neurotoxicity of natural β-amyloid peptides when contacted with the natural β-amyloid peptides. In a preferred embodiment, Xaa_1 and Xaa_2 are each phenylalanine structures. In another preferred embodiment Xaa_2 and Xaa_3 are each phenylalanine structures.

The invention further provides a β-amyloid modulator compound comprising a formula:



wherein Xaa_1 and Xaa_3 are amino acid structures;

Xaa_2 is a valine structure;

Xaa_4 is a phenylalanine structure;

Y, which may or may not be present, is a peptidic structure having the formula $(Xaa)_a$, wherein Xaa is any amino acid structure and a is an integer from 1 to 15;

Z, which may or may not be present, is a peptidic structure having the formula $(Xaa)_b$, wherein Xaa is any amino acid structure and b is an integer from 1 to 15; and

A is a modifying group attached directly or indirectly to the compound and n is an integer;

Xaa_1, Xaa_3, Y, Z, A and n being selected such that the compound modulates the aggregation or inhibits the neurotoxicity of natural β-amyloid peptides when contacted with the natural β-amyloid peptides. In a preferred embodiment, Xaa_1 is a leucine structure and Xaa_3 is phenylalanine structure.

The invention still further provides a compound comprising the formula:

$$A\text{-}Xaa_1\text{-}Xaa_2\text{-}Xaa_3\text{-}Xaa_4\text{-}Xaa_5\text{-}Xaa_6\text{-}Xaa_7\text{-}Xaa_8\text{-}B$$

wherein Xaa1 is a histidine structure;

Xaa2 is a glutamine structure;

Xaa3 is a lysine structure;

Xaa4 is a leucine structure;

Xaa5 is a valine structure;

Xaa6 is a phenylalanine structure;

Xaa7 is a phenylalanine structure;

Xaa8 is an alanine structure;

A and B are modifying groups attached directly or indirectly to the amino terminus and carboxy terminus, respectively, of the compound;

and wherein Xaa_1-Xaa_2-Xaa_3, Xaa_1-Xaa_2 or Xaa_1 may or may not be present;

Xaa_8 may or may not be present; and

at least one of A and B is present.

The invention still further provides a β-amyloid modulator compound comprising a modifying group attached directly or indirectly to a peptidic structure, wherein the peptidic structure comprises amino acid structures having an amino acid sequence selected from the group consisting of His-Gln-Lys-Leu-Val-Phe-Phe-Ala (SEQ ID NO:5), His-Gln-Lys-Leu-Val-Phe-Phe (SEQ ID NO:6), Gln-Lys-Leu-Val-Phe-Phe-Ala (SEQ ID NO:7), Gln-Lys-Leu-Val-Phe-Phe (SEQ ID NO:8), Lys-Leu-Val-Phe-Phe-Ala (SEQ ID NO:9), Lys-Leu-Val-Phe-Phe (SEQ ID NO:10), Leu-Val-

Phe-Phe-Ala (SEQ ID NO:11), Leu-Val-Phe-Phe (SEQ ID NO:12), Leu-Ala-Phe-Phe-Ala (SEQ ID NO:13), Val-Phe-Phe (SEQ ID NO:19), Phe-Phe-Ala (SEQ ID NO:20), Phe-Phe-Val-Leu-Ala (SEQ ID NO:21), Leu-Val-Phe-Phe-Lys (SEQ ID NO:22), Leu-Val-Iodotyrosine-Phe-Ala (SEQ ID NO:23), Val-Phe-Phe-Ala (SEQ ID NO:24), Ala-Val-Phe-Phe-Ala (SEQ ID NO:25), Leu-Val-Phe-Iodotyrosine-Ala (SEQ ID NO:26), Leu-Val-Phe-Phe-Ala-Glu (SEQ ID NO:27), Phe-Phe-Val-Leu (SEQ ID NO:28), Phe-Lys-Phe-Val-Leu (SEQ ID NO:29), Lys-Leu-Val-Ala-Phe (SEQ ID NO:30), Lys-Leu-Val-Phe-Phe-βAla (SEQ ID NO:31) and Leu-Val-Phe-Phe-DAla (SEQ ID NO:32).

In the compounds of the invention comprising a modifying group, preferably the modifying group comprises a cyclic, heterocyclic or polycyclic group. Preferred modifying groups contains a cis-decalin group, such as a cholanoyl structure. Preferred modifying groups include a cholyl group, a biotin-containing group, a diethylenetriaminepentaacetyl group, a (−)-menthoxyacetyl group, a fluorescein-containing group or an N-acetylneuraminyl group.

The compounds of the invention can be further modified, for example to alter a pharmacokinetic property of the compound or to label the compound with a detectable substance. Preferred radioactive labels are radioactive iodine or technetium.

The invention also provides a β-amyloid modulator which inhibits aggregation of natural β-amyloid peptides when contacted with a molar excess amount of natural β-amyloid peptides.

The invention also provides a β-amyloid peptide compound comprising an amino acid sequence having at least one amino acid deletion compared to βAP_{1-39}, such that the compound inhibits aggregation of natural β-amyloid peptides when contacted with the natural β-amyloid peptides. In one embodiment, the compound has at least one internal amino acid deleted compared to βAP_{1-39}. In another embodiment, the compound has at least one N-terminal amino acid deleted compared to βAP_{1-39}. In yet another embodiment, the compound has at least one C-terminal amino acid deleted compared to βAP_{1-39}. Preferred compounds include βAP_{6-20} (SEQ ID NO:4), βAP_{16-30} (SEQ ID NO:14), $\beta AP_{1-20,\ 26-40}$ (SEQ ID NO:15) and EEVVHHHHQQ-βAP_{16-40} (SEQ ID NO:16).

The compounds of the invention can be formulated into pharmaceutical compositions comprising the compound and a pharmaceutically acceptable carrier. The compounds can also be used in the manufacture of a medicament for the diagnosis or treatment of an amyloidogenic disease.

Another aspect of the invention pertains to diagnostic and treatment methods using the compounds of the invention. The invention provides a method for inhibiting aggregation of natural β-amyloid peptides, comprising contacting the natural β-amyloid peptides with a compound of the invention such that aggregation of the natural β-amyloid peptides is inhibited. The invention also provides a method for inhibiting neurotoxicity of natural β-amyloid peptides, comprising contacting the natural β-amyloid peptides with a compound of the invention such that neurotoxicity of the natural β-amyloid peptides is inhibited.

In another embodiment, the invention provides a method for detecting the presence or absence of natural β-amyloid peptides in a biological sample, comprising contacting a biological sample with a compound of the invention and detecting the compound bound to natural β-amyloid peptides to thereby detect the presence or absence of natural β-amyloid peptides in the biological sample. In one embodiment, the β-amyloid modulator compound and the biological sample are contacted in vitro. In another embodiment, the β-amyloid modulator compound is contacted with the biological sample by administering the β-amyloid modulator compound to a subject. For in vivo administration, preferably the compound is labeled with radioactive technetium or radioactive iodine.

In another embodiment, the invention provides a method for detecting natural β-amyloid peptides to facilitate diagnosis of a β-amyloidogenic disease, comprising contacting a biological sample with a compound of the invention and detecting the compound bound to natural β-amyloid peptides to facilitate diagnosis of a β-amyloidogenic disease. In one embodiment, the β-amyloid modulator compound and the biological sample are contacted in vitro. In another embodiment, the β-amyloid modulator compound is contacted with the biological sample by administering the β-amyloid modulator compound to a subject. For in vivo administration, preferably the compound is labeled with radioactive technetium or radioactive iodine. Preferably, the method facilitates diagnosis of Alzheimer's disease.

The invention also provides a method for treating a subject for a disorder associated with amyloidosis, comprising administering to the subject a therapeutically or prophylactically effective amount of a compound of the invention such that the subject is treated for a disorder associated with amyloidosis. The method can be used to treat disorders is selected, for example, from the group consisting of familial amyloid polyneuropathy (Portuguese, Japanese and Swedish types), familial amyloid cardiomyopathy (Danish type), isolated cardiac amyloid, systemic senile amyloidosis, scrapie, bovine spongiform encephalopathy, Creutzfeldt-Jakob disease, Gerstmann-Straussler-Scheinker syndrome, adult onset diabetes, insulinoma, isolated atrial amyloidosis, idiopathic (primary) amyloidosis, myeloma or macroglobulinemia-associated amyloidosis, primary localized cutaneous nodular amyloidosis associated with Sjogren's syndrome, reactive (secondary) amyloidosis, familial Mediterranean Fever and familial amyloid nephropathy with urticaria and deafness (Muckle-Wells syndrome), hereditary cerebral hemorrhage with amyloidosis of Icelandic type, amyloidosis associated with long term hemodialysis, hereditary non-neuropathic systemic amyloidosis (familial amyloid polyneuropathy III), familial amyloidosis of Finnish type, amyloidosis associated with medullary carcinoma of the thyroid, fibrinogen-associated hereditary renal amyloidosis and lysozyme-associated hereditary systemic amyloidosis.

In a preferred embodiment, the invention provides a method for treating a subject for a disorder associated with β-amyloidosis, comprising administering to the subject a therapeutically or prophylactically effective amount of a compound of the invention such that the subject is treated for a disorder associated with β-amyloidosis. Preferably the disorder is Alzheimer's disease.

In yet another embodiment, the invention provides a method for treating a subject for a disorder associated with β-amyloidosis, comprising administering to the subject a recombinant expression vector encoding a peptide compound of the invention such that the compound is synthesized in the subject and the subject is treated for a disorder associated with β-amyloidosis. Preferably, the disorder is Alzheimer's disease.

BRIEF DESCRIPTION OF THE DRAWING

FIG. 1 is a graphic representation of the turbidity of a β-AP_{1-40} solution, as measured by optical density at 400 nm,

either in the absence of a β-amyloid modulator or in the presence of the β-amyloid modulator N-biotinyl-βAP_{1-40} (1%, or 5%).

FIG. 2 is a schematic representation of compounds which can be used to modify a β-AP or an Aβ aggregation core domain to form a β-amyloid modulator of the invention.

FIG. 3 is a graphic representation of the toxicity of $Aβ_{1-40}$ aggregates, but not $Aβ_{1-40}$ monomers, to cultured neuronal cells.

FIGS. 4A–F are a graphic representation of the aggregation of $Aβ_{1-40}$ in the presence of an equimolar amount of cholyl-$Aβ_{6-20}$ (panel A), a ~2-fold molar excess of cholyl-$Aβ_{6-20}$ (panel B) or a ~6-fold molar excess of cholyl-$Aβ_{6-20}$ (panel C) and the corresponding toxicity of the aggregates of panels A, B and C to cultured neuronal cells (panels D, E and F, respectively).

DETAILED DESCRIPTION OF THE INVENTION

This invention pertains to compounds, and pharmaceutical compositions thereof, that can modulate the aggregation of amyloidogenic proteins and peptides, in particular compounds that can modulate the aggregation of natural D amyloid peptides (β-AP) and inhibit the neurotoxicity of natural β-APs. A compound of the invention that modulates aggregation of natural β-AP, referred to herein interchangeably as a β amyloid modulator compound, a β amyloid modulator or simply a modulator, alters the aggregation of natural β-AP when the modulator is contacted with natural β-AP. Thus, a compound of the invention acts to alter the natural aggregation process or rate for β-AP, thereby disrupting this process. Preferably, the compounds inhibit β-AP aggregation. Furthermore, the invention provides subregions of the β amyloid peptide that are sufficient, when appropriately modified as described herein, to alter (and preferably inhibit) aggregation of natural β amyloid peptides when contacted with the natural β amyloid peptides. In particular, preferred modulator compounds of the invention are comprised of a modified form of an Aβ aggregation core domain, modeled after the aforementioned Aβ subregion (as described further below), which is sufficient to alter (and preferably inhibit) the natural aggregation process or rate for β-AP. This Aβ aggregation core domain can comprises as few as three amino acid residues (or derivative, analogues or mimetics thereof). Moreover, while the amino acid sequence of the Aβ aggregation core domain can directly correspond to an amino acid sequence found in natural β-AP, it is not essential that the amino acid sequence directly correspond to a β-AP sequence. Rather, amino acid residues derived from a preferred subregion of β-AP (a hydrophobic region centered around positions 17–20) can be rearranged in order and/or substituted with homologous residues within a modulator compound of the invention and yet maintain their inhibitory activity (described further below).

The β amyloid modulator compounds of the invention can be selected based upon their ability to inhibit the aggregation of natural β-AP in vitro and/or inhibit the neurotoxicity of natural β-AP fibrils for cultured cells (using assays described herein). Accordingly, the preferred modulator compounds inhibit the aggregation of natural β-AP and/or inhibit the neurotoxicity of natural β-AP. However, modulator compounds selected based on one or both of these properties may have additional properties in vivo that may be beneficial in the treatment of amyloidosis. For example, the modulator compound may interfere with processing of natural β-AP (either by direct or indirect protease inhibition) or by modulation of processes that produce toxic β-AP, or other APP fragments, in vivo. Alternatively, modulator compounds may be selected based on these latter properties, rather than inhibition of Aβ aggregation in vitro. Moreover, modulator compounds of the invention that are selected based upon their interaction with natural β-AP also may interact with APP or with other APP fragments.

As used herein, a "modulator" of β-amyloid aggregation is intended to refer to an agent that, when contacted with natural β amyloid peptides, alters the aggregation of the natural β amyloid peptides. The term "aggregation of β amyloid peptides" refers to a process whereby the peptides associate with each other to form a multimeric, largely insoluble complex. The term "aggregation" further is intended to encompass β amyloid fibril formation and also encompasses β-amyloid plaques.

The terms "natural β-amyloid peptide", "natural β-AP" and "natural Aβ peptide", used interchangeably herein, are intended to encompass naturally occurring proteolytic cleavage products of the β amyloid precursor protein (APP) which are involved in β-AP aggregation and β-amyloidosis. These natural peptides include β-amyloid peptides having 39–43 amino acids (i.e., $Aβ_{1-39}$, $Aβ_{1-40}$, $Aβ_{1-41}$, $Aβ_{1-42}$ and $Aβ_{1-43}$). The amino-terminal amino acid residue of natural β-AP corresponds to the aspartic acid residue at position 672 of the 770 amino acid residue form of the amyloid precursor protein ("APP-770"). The 43 amino acid long form of natural β-AP has the amino acid sequence.

DAEFRHDSGYEVHHQKLVFFAEDVGSNK-GAIIGLMVGGVVIAT

(also shown in SEQ ID NO:1), whereas the shorter forms have 1–4 amino acid residues deleted from the carboxy-terminal end. The amino acid sequence of APP-770 from position 672 (i.e., the amino-terminus of natural β-AP) to its C-terminal end (103 amino acids) is shown in SEQ ID NO:2. The preferred form of natural β-AP for use in the aggregation assays described herein is $Aβ_{1-40}$.

In the presence of a modulator of the invention, aggregation of natural β amyloid peptides is "altered" or "modulated". The various forms of the term "alteration" or "modulation" are intended to encompass both inhibition of β-AP aggregation and promotion of β-AP aggregation. Aggregation of natural β-AP is "inhibited" in the presence of the modulator when there is a decrease in the amount and/or rate of β-AP aggregation as compared to the amount and/or rate of β-AP aggregation in the absence of the modulator. The various forms of the term "inhibition" are intended to include both complete and partial inhibition of β-AP aggregation. Inhibition of aggregation can be quantitated as the fold increase in the lag time for aggregation or as the decrease in the overall plateau level of aggregation (i.e., total amount of aggregation), using an aggregation assay as described in the Examples. In various embodiments, a modulator of the invention increases the lag time of aggregation at least 1.2-fold, 1.5-fold, 1.8-fold, 2-fold, 2.5-fold, 3-fold, 4-fold or 5-fold. In various other embodiments, a modulator of the invention inhibits the plateau level of aggregation at least 10%, 20%, 30%, 40%, 50%, 75% or 100%.

A modulator which inhibits β-AP aggregation (an "inhibitory modulator compound") can be used to prevent or delay the onset of β-amyloid deposition. Moreover, as demonstrated in Example 10, inhibitory modulator compounds of the invention inhibit the formation and/or activity of neurotoxic aggregates of natural Aβ peptide (i.e., the inhibitory compounds can be used to inhibit the neurotoxicity of

β-AP). Still further, also as demonstrated in Example 10, the inhibitory compounds of the invention can be used to reduce the neurotoxicity of preformed β-AP aggregates, indicating that the inhibitory modulators can either bind to preformed Aβ fibrils or soluble aggregate and modulate their inherent neurotoxicity or that the modulators can perturb the equilibrium between monomeric and aggregated forms of β-AP in favor of the non-neurotoxic form.

Alternatively, in another embodiment, a modulator compound of the invention promotes the aggregation of natural Aβ peptides. The various forms of the term "promotion" refer to an increase in the amount and/or rate of β-AP aggregation in the presence of the modulator, as compared to the amount and/or rate of β-AP aggregation in the absence of the modulator. Such a compound which promotes Aβ aggregation is referred to as a stimulatory modulator compound. Stimulatory modulator compounds may be useful for sequestering β-amyloid peptides, for example in a biological compartment where aggregation of β-AP may not be deleterious to thereby deplete β-AP from a biological compartment where aggregation of β-AP is deleterious. Moreover, stimulatory modulator compounds can be used to promote Aβ aggregation in in vitro aggregation assays (e.g., assays such as those described in the Examples), for example in screening assays for test compounds that can then inhibit or reverse this Aβ aggregation (i.e., a stimulatory modulator compound can act as a "seed" to promote the formation of Aβ aggregates).

In a preferred embodiment, the modulators of the invention are capable of altering β-AP aggregation when contacted with a molar excess amount of natural β-AP. A "molar excess amount of natural β-AP" refers to a concentration of natural β-AP, in moles, that is greater than the concentration, in moles, of the modulator. For example, if the modulator and β-AP are both present at a concentration of 1 μM, they are said to be "equimolar", whereas if the modulator is present at a concentration of 1 μM and the β-AP is present at a concentration of 5 μM, the β-AP is said to be present at a 5-fold molar excess amount compared to the modulator. In preferred embodiments, a modulator of the invention is effective at altering natural β-AP aggregation when the natural β-AP is present at at least a 2-fold, 3-fold or 5-fold molar excess compared to the concentration of the modulator. In other embodiments, the modulator is effective at altering β-AP aggregation when the natural β-AP is present at at least a 10-fold, 20-fold, 33-fold, 50-fold, 100-fold, 500-fold or 1000-fold molar excess compared to the concentration of the modulator.

Various additional aspects of the modulators of the invention, and the uses thereof, are described in further detail in the following subsections.

I. Modulator Compounds

In one embodiment, a modulator of the invention comprises a β-amyloid peptide compound comprising the formula:

wherein Xaa is a β-amyloid peptide, A is a modulating group attached directly or indirectly to the β-amyloid peptide of the compound such that the compound inhibits aggregation of natural β-amyloid peptides when contacted with the natural β-amyloid peptides, and n is an integer selected such that the compound inhibits aggregation of natural β-amyloid peptides when contacted with the natural β-amyloid peptides.

Preferably, β-amyloid peptide of the compound has an amino-terminal amino acid residue corresponding to position 668 of β-amyloid precursor protein-770 (APP-770) or to a residue carboxy-terminal to position 668 of APP-770. The amino acid sequence of APP-770 from position 668 to position 770 (i.e., the carboxy terminus) is shown below and in SEQ ID NO:2:

EVKMDAEFRHDSGYEVHHQKLVFFAED-VGSNKGAIIGLMVGGVVIATVIVITLVM-LKKKQYTSIHHGVVEVDAAVTPEERHL-SKMQQNGYENPTYKFFEQMQN.

More preferably, the amino-terminal amino acid residue of the β-amyloid peptide corresponds to position 672 of APP-770 (position 5 of the amino acid sequence of SEQ ID NO:2) or to a residue carboxy-terminal to position 672 of APP-770. Although the β-amyloid peptide of the compound may encompass the 103 amino acid residues corresponding to positions 668–770 of APP-770, preferably the peptide is between 6 and 60 amino acids in length, more preferably between 10 and 43 amino acids in length and even more preferably between 10 and 25 amino acid residues in length.

As used herein, the term "β amyloid peptide", as used in a modulator of the invention is intended to encompass peptides having an amino acid sequence identical to that of the natural sequence in APP, as well as peptides having acceptable amino acid substitutions from the natural sequence. Acceptable amino acid substitutions are those that do not affect the ability of the peptide to alter natural β-AP aggregation. Moreover, particular amino acid substitutions may further contribute to the ability of the peptide to alter natural β-AP aggregation and/or may confer additional beneficial properties on the peptide (e.g., increased solubility, reduced association with other amyloid proteins, etc.). For example, substitution of hydrophobic amino acid residues for the two phenylalanine residues at positions 19 and 20 of natural β-AP (positions 19 and 20 of the amino acid sequence shown in SEQ ID NO:1) may further contribute to the ability of the peptide to alter β-AP aggregation (see Hilbich, C. (1992) *J. Mol. Biol.* 228:460–473). Thus, in one embodiment, the β-AP of the compound consists of the amino acid sequence shown below and in SEQ ID NO:3:

DAEFRHDSGYEVHHQKLV(Xaa_{19})(Xaa_{20})AEDVGSNKGAIIGLMVGGVVIAT

(or an amino-terminal or carboxy-terminal deletion thereof), wherein Xaa is a hydrophobic amino acid. Examples of hydrophobic amino acids are isoleucine, leucine, threonine, serine, alanine, valine or glycine. Preferably, $F_{19}F_{20}$ is substituted with $T_{19}T_{20}$ or $G_{19}I_{20}$.

Other suitable amino acid substitutions include replacement of amino acids in the human peptide with the corresponding amino acids of the rodent β-AP peptide. The three amino acid residues that differ between human and rat β-AP are at positions 5, 10 and 13 of the amino acid sequence shown in SEQ ID NOs:1 and 3. A human β-AP having the human to rodent substitutions Arg_5 to Gly, Tyr_{10} to Phe and His_{13} to Arg has been shown to retain the properties of the human peptide (see Fraser, P. E. et al. (1992) *Biochemistry* 31:10716–10723; and Hilbich, C. et al. (1991) *Eur. J. Biochem.* 201:61–69). Accordingly, a human β-AP having rodent β-AP a.a. substitutions is suitable for use in a modulator of the invention.

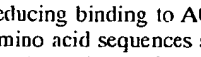
Other possible β-AP amino acid substitutions are described in Hilbich, C. et al. (1991) *J. Mol. Biol.* 218:149–163; and Hilbich, C. (1992) *J. Mol. Biol.* 228:460–473. Moreover, amino acid substitutions that affect the ability of β-AP to associate with other proteins can be introduced. For example, one or more amino acid substitutions that reduce the ability of β-AP to associate with the serpin enzyme complex (SEC) receptor,

α1-antichymotrypsin (ACT) and/or apolipoprotein E (ApoE) can be introduced. A preferred substitution for reducing binding to the SEC receptor is $L_{34}M_{35}$ to $A_{34}A_{35}$ (at positions 34 and 35 of the amino acid sequences shown in SEQ ID NOs:1 and 3). A preferred substitution for reducing binding to ACT is S_8 to A_8 (at position 8 of the amino acid sequences shown in SEQ ID NOs:1 and 3).

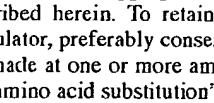
Alternative to β-AP amino acid substitutions described herein or known in the art, a modulator composed, at least in part, of an amino acid-substituted β amyloid peptide can be prepared by standard techniques and tested for the ability to alter β-AP aggregation using an aggregation assay described herein. To retain the properties of the original modulator, preferably conservative amino acid substitutions are made at one or more amino acid residues. A "conservative amino acid substitution" is one in which the amino acid residue is replaced with an amino acid residue having a similar side chain. Families of amino acid residues having similar side chains have been defined in the art, including basic side chains (e.g., lysine, arginine, histidine), acidic side chains (e.g., aspartic acid, glutamic acid), uncharged polar side chains (e.g. glycine, asparagine, glutamine, serine, threonine, tyrosine, cysteine), nonpolar side chains (e.g., alanine, valine, leucine, isoleucine, proline, phenylalanine, methionine, tryptophan), β-branched side chains (e.g., threonine, valine, isoleucine) and aromatic side chains (e.g., tyrosine, phenylalanine, tryptophan, histidine). Accordingly, a modulator composed of a β amyloid peptide having an amino acid sequence that is mutated from that of the wild-type sequence in APP-770 yet which still retains the ability to alter natural β-AP aggregation is within the scope of the invention.

As used herein, the term "β amyloid peptide" is further intended to include peptide analogues or peptide derivatives or peptidomimetics that retain the ability to alter natural β-AP aggregation as described herein. For example, a β amyloid peptide of a modulator of the invention may be modified to increase its stability, bioavailability, solubility, etc. The terms "peptide analogue", "peptide derivative" and "peptidomimetic" as sued herein are intended to include molecules which mimic the chemical structure of a peptide and retain the functional properties of the peptide. Approaches to designing peptide analogs are known in the art. For example, see Farmer, P. S. in *Drug Design* (E. J. Ariens, ed.) Academic Press, New York, 1980, vol. 10, pp. 119–143; Ball. J. B. and Alewood, P. F. (1990) *J. Mol. Recognition* 3:55; Morgan, B. A. and Gainor, J. A. (1989) *Ann. Rep. Med. Chem.* 24:243; and Freidinger, R. M. (1989) *Trends Pharmacol. Sci.* 10:270. Examples of peptide analogues, derivatives and peptidomimetics include peptides substituted with one or more benzodiazepine molecules (see e.g., James, G. L. et al. (1993) *Science* 260:1937–1942), peptides with methylated amide linkages and "retro-inverso" peptides (see U.S. Pat. No. 4,522,752 by Sisto). Peptide analogues, peptide derivatives and peptidomimetic are described in further detail below with regard to compounds comprising an Aβ aggregation core domain.

In a modulator of the invention having the formula shown above, a modulating group ("A") is attached directly or indirectly to the β-amyloid peptide of the modulator (As used herein, the term "modulating group" and "modifying group" are used interchangeably to describe a chemical group directly or indirectly attached to an Aβ derived peptidic structure). For example, the modulating group can be directly attached by covalent coupling to the β-amyloid peptide or the modulating group can be attached indirectly by a stable non-covalent association. In one embodiment of the invention, the modulating group is attached to the amino-terminus of the β-amyloid peptide of the modulator. Accordingly, the modulator can comprise a compound having a formula:

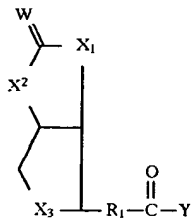


Alternatively, in another embodiment of the invention, the modulating group is attached to the carboxy-terminus of the β-amyloid peptide of the modulator. Accordingly, the modulator can comprise a compound having a formula:



In yet another embodiment, the modulating group is attached to the side chain of at least one amino acid residue of the β-amyloid peptide of the compound (e.g., through the epsilon amino group of a lysyl residue(s), through the carboxyl group of an aspartic acid residue(s) or a glutamic acid residue(s), through a hydroxy group of a tyrosyl residue (s), a serine residue(s) or a threonine residue(s) or other suitable reactive group on an amino acid side chain).

The modulating group is selected such that the compound inhibits aggregation of natural β-amyloid peptides when contacted with the natural β-amyloid peptides. Accordingly, since the β-AP peptide of the compound is modified from its natural state, the modulating group "A" as used herein is not intended to include hydrogen. In a preferred embodiment, the modulating group is a biotin compound of the formula:



wherein X_1–X_3 are each independently selected from the group consisting of S, O and NR_2, wherein R_2 is hydrogen, or an aryl, lower alkyl, alkenyl or alkynyl moiety; W is O or NR_2; R_1 is a lower alkylenyl moiety and Y is a direct bond or a spacer molecule selected for its ability to react with a target group on a β-AP. At least one of X_1–X_3 is an NR_2 group or W is an $N(R_2)_2$ group.

The term "aryl" is intended to include aromatic moieties containing substituted or unsubstituted ring(s), e.g., benzyl, naphthyl, etc. Other more complex fused ring moieties also are intended to be included.

The term "lower alkyl or alkylenyl moiety" refers to a saturated, straight or branched chain (or combination thereof) hydrocarbon containing 1 to about 6 carbon atoms, more preferably from 1 to 3 carbon atoms. The terms "lower alkenyl moiety" and "lower alkynyl moiety" refer to unsaturated hydrocarbons containing 1 to about 6 carbon atoms, more preferably 1 to 3 carbon atoms. Preferably, R_2 contains 1 to 3 carbon atoms. Preferably, R_1 contains 4 carbon atoms.

The spacer molecule (Y) can be, for example, a lower alkyl group or a linker peptide, and is preferably selected for its ability to link with a free amino group (e.g., the α-amino group at the amino-terminus of a β-AP). Thus, in a preferred embodiment, the biotin compound modifies the amino-terminus of a β-amyloid peptide.

Additional suitable modulating groups may include other cyclic and heterocyclic compounds and other compounds having similar steric "bulk". Non-limiting examples of compounds which can be used to modify a β-AP are shown schematically in FIG. 2, and include N-acetylneuraminic acid, cholic acid, trans-4-cotininecarboxylic acid, 2-imino-1-imidazolidineacetic acid, (S)-(−)-indoline-2-carboxylic acid, (−)-menthoxyacetic acid, 2-norbornaneacetic acid, γ-oxo-5-acenaphthenebutyric acid, (−)-2-oxo-4-thiazolidinecarboxylic acid, tetrahydro-3-furoic acid, 2-iminobiotin-N-hydroxysuccinimide ester, diethylenetriaminepentaacetic dianhydride, 4-morpholinecarbonyl chloride, 2-thiopheneacetyl chloride, 2-thiophenesulfonyl chloride, 5-(and 6-)-carboxyfluorescein (succinimidyl ester), fluorescein isothiocyanate, and acetic acid (or derivatives thereof). Suitable modulating groups are described further in subsection II below.

In a modulator of the invention, a single modulating group may be attached to a β-amyloid peptide (e.g., n=1 in the formula shown above) or multiple modulating groups may be attached to the peptide. The number of modulating groups is selected such that the compound inhibits aggregation of natural β-amyloid peptides when contacted with the natural β-amyloid peptides. However, n preferably is an integer between 1 and 60, more preferably between 1 and 30 and even more preferably between 1 and 10 or 1 and 5.

In another embodiment, a β-amyloid modulator compound of the invention comprises an Aβ aggregation core domain (abbreviated as ACD) coupled directly or indirectly to a modifying group such that the compound modulates the aggregation or inhibits the neurotoxicity of natural β-amyloid peptides when contacted with the natural β-amyloid peptides. As used herein, an "Aβ aggregation core domain" is intended to refer to a structure that is modeled after a subregion of a natural β-amyloid peptide which is sufficient to modulate aggregation of natural β-APs when this subregion of the natural β-AP is appropriately modified as described herein (e.g., modified at the amino-terminus). The term "subregion of a natural β-amyloid peptide" is intended to include amino-terminal and/or carboxy-terminal deletions of natural β-AP. The term "subregion of natural β-AP" is not intended to include full-length natural β-AP (i.e., "subregion" does not include $A\beta_{1-39}$, $A\beta_{1-40}$, $A\beta_{1-41}$, $A\beta_{1-42}$ and $A\beta_{1-43}$).

Although not intending to be limited by mechanism, the ACD of the modulators of the invention is thought to confer a specific targeting function on the compound that allows the compound to recognize and specifically interact with natural β-AP. Preferably, the ACD is modeled after a subregion of natural β-AP that is less than 15 amino acids in length and more preferably is between 3–10 amino acids in length. In various embodiments, the ACD is modeled after a subregion of β-AP that is 10, 9, 8, 7, 6, 5, 4 or 3 amino acids in length. In one embodiment, the subregion of β-AP upon which the ACD is modeled is an internal or carboxy-terminal region of β-AP (i.e., downstream of the amino-terminus at amino acid position 1). In another embodiment, the ACD is modeled after a subregion of β-AP that is hydrophobic. In certain specific embodiments, the term Aβ aggregation core domain specifically excludes β-AP subregions corresponding to amino acid positions 1–15 ($A\beta_{1-15}$), 6–20 ($A\beta_{6-20}$) and 16–40 ($A\beta_{16-40}$).

An Aβ aggregation core domain can be comprised of amino acid residues linked by peptide bonds. That is, the ACD can be a peptide corresponding to a subregion of β-AP. Alternatively, an Aβ aggregation core domain can be modeled after the natural Aβ peptide region but may be comprised of a peptide analogue, peptide derivative or peptidomimetic compound, or other similar compounds which mimics the structure and function of the natural peptide. Accordingly, as used herein, an "Aβ aggregation core domain" is intended to include peptides, peptide analogues, peptide derivatives and peptidomimetic compounds which, when appropriately modified, retain the aggregation modulatory activity of the modified natural Aβ peptide subregion. Such structures that are designed based upon the amino acid sequence are referred to herein as "Aβ derived peptidic structures." Approaches to designing peptide analogues, derivatives and mimetics are known in the art. For example, see Farmer, P. S. in *Drug Design* (E. J. Ariens, ed.) Academic Press, New York, 1980, vol. 10, pp. 119–143; Ball. J. B. and Alewood, P. F. (1990) *J. Mol. Recognition* 3:55; Morgan, B. A. and Gainor, J. A. (1989) *Ann. Rep. Med Chem.* 24:243; and Freidinger, R. M. (1989) *Trends Pharmacol. Sci.* 10:270. See also Sawyer, T. K. (1995) "Peptidomimetic Design and Chemical Approaches to Peptide Metabolism" in Taylor, M. D. and Amidon, G. L. (eds.) *Peptide-Based Drug Design: Controlling Transport and Metabolism*, Chapter 17; Smith, A. B. 3rd, et al. (1995) *J. Am. Chem. Soc.* 117:11113–11123; Smith, A. B. 3rd, et al. (1994) *J. Am. Chem. Soc.* 116:9947–9962; and Hirschman, R., et al. (1993) *J. Am. Chem. Soc.* 115:12550–12568.

As used herein, a "derivative" of a compound X (e.g., a peptide or amino acid) refers to a form of X in which one or more reaction groups on the compound have been derivatized with a substituent group. Examples of peptide derivatives include peptides in which an amino acid side chain, the peptide backbone, or the amino- or carboxy-terminus has been derivatized (e.g., peptidic compounds with methylated amide linkages). As used herein an "analogue" of a compound X refers to a compound which retains chemical structures of X necessary for functional activity of X yet which also contains certain chemical structures which differ from X. An examples of an analogue of a naturally-occurring peptide is a peptides which includes one or more non-naturally-occurring amino acids. As used herein, a "mimetic" of a compound X refers to a compound in which chemical structures of X necessary for functional activity of X have been replaced with other chemical structures which mimic the conformation of X. Examples of peptidomimetics include peptidic compounds in which the peptide backbone is substituted with one or more benzodiazepine molecules (see e.g., James, G. L. et al. (1993) *Science* 260:1937–1942), peptides in which all L-amino acids are substituted with the corresponding D-amino acids and "retro-inverso" peptides (see U.S. Pat. No. 4,522,752 by Sisto), described further below.

The term mimetic, and in particular, peptidomimetic, is intended to include isosteres. The term "isostere" as used herein is intended to include a chemical structure that can be substituted for a second chemical structure because the steric conformation of the first structure fits a binding site specific for the second structure. The term specifically includes peptide back-bone modifications (i.e., amide bond mimetics) well known to those skilled in the art. Such modifications include modifications of the amide nitrogen, the α-carbon, amide carbonyl, complete replacement of the amide bond, extensions, deletions or backbone crosslinks. Several peptide backbone modifications are known, including $\psi[CH_2S]$, $\psi[CH_2NH]$, $\psi[CSNH_2]$, $\psi[NHCO]$, $\psi[COCH_2]$, and $\psi[(E)$ or (Z) $CH{=}CH]$. In the nomenclature used above, ψ indicates the absence of an amide bond. The structure that replaces the amide group is specified within the brackets. Other examples of isosteres include peptides substituted

with one or more benzodiazepine molecules (see e.g., James, G. L. et al. (1993) *Science* 260:1937–1942).

Other possible modifications include an N-alkyl (or aryl) substitution (ψ[CONR]), backbone crosslinking to construct lactams and other cyclic structures, substitution of all D-amino acids for all L-amino acids within the compound ("inverso" compounds) or retro-inverso amino acid incorporation (ψ[NHCO]). By "inverso" is meant replacing L-amino acids of a sequence with D-amino acids, and by "retro-inverso" or "enantio-retro" is meant reversing the sequence of the amino acids ("retro") and replacing the L-amino acids with D-amino acids. For example, if the parent peptide is Thr-Ala-Tyr, the retro modified form is Tyr-Ala-Thr, the inverso form is thr-ala-tyr, and the retro-inverso form is tyr-ala-thr (lower case letters refer to D-amino acids). Compared to the parent peptide, a retro-inverso peptide has a reversed backbone while retaining substantially the original spatial conformation of the side chains, resulting in a retro-inverso isomer with a topology that closely resembles the parent peptide. See Goodman et al. "Perspectives in Peptide Chemistry" pp. 283–294 (1981). See also U.S. Pat. No. 4,522,752 by Sisto for further description of "retro-inverso" peptides.

Other derivatives of the modulator compounds of the invention include C-terminal hydroxymethyl derivatives, O-modified derivatives (e.g., C-terminal hydroxymethyl benzyl ether), N-terminally modified derivatives including substituted amides such as alkylamides and hydrazides and compounds in which a C-terminal phenylalanine residue is replaced with a phenethylamide analogue (e.g., Val-Phe-phenethylamide as an analogue of the tripeptide Val-Phe-Phe).

In a preferred embodiment, the ACD of the modulator is modeled after the subregion of β-AP encompassing amino acid positions 17–20 (i.e., Leu-Val-Phe-Phe; SEQ ID NO:12). As described further in Examples 7, 8 and 9, peptide subregions of $A\beta_{1-40}$ were prepared, amino-terminally modified and evaluated for their ability to modulate aggregation of natural β-amyloid peptides. One subregion that was effective at inhibiting aggregation was $A\beta_{6-20}$ (i.e., amino acid residues 6–20 of the natural $A\beta_{1-40}$ peptide, the amino acid sequence of which is shown in SEQ ID NO:4). Amino acid residues were serially deleted from the amino-terminus or carboxy terminus of this subregion to further delineate a minimal subregion that was sufficient for aggregation inhibitory activity. This process defined $A\beta_{17-20}$ (i.e., amino acid residues 17–20 of the natural $A\beta_{1-40}$ peptide) as a minimal subregion that, when appropriately modified, is sufficient for aggregation inhibitory activity. Accordingly, an "Aβ aggregation core domain" within a modulator compound of the invention can be modeled after $A\beta_{17-20}$. In one embodiment, the Aβ aggregation core domain comprises $A\beta_{17-20}$ itself (i.e., a peptide comprising the amino acid sequence leucine-valine-phenylalanine-phenylalanine; SEQ ID NO:12). In other embodiments, the structure of $A\beta_{17-20}$ is used as a model to design an Aβ aggregation core domain having similar structure and function to $A\beta_{17-20}$. For example, peptidomimetics, derivatives or analogues of $A\beta_{17-20}$ (as described above) can be used as an Aβ aggregation core domain. In addition to $A\beta_{17-20}$, the natural Aβ peptide is likely to contain other minimal subregions that are sufficient for aggregation inhibitory activity. Such additional minimal subregions can be identified by the processes described in Examples 7, 8 and 9, wherein a 15 mer subregion of $A\beta_{1-40}$ is serially deleted from the amino-terminus or carboxy terminus, the deleted peptides are appropriately modified and then evaluated for aggregation inhibitory activity.

One form of the β-amyloid modulator compound comprising an Aβ aggregation core domain modeled after $A\beta_{17-20}$ coupled directly or indirectly to at least one modifying group has the formula:

$$(Y—Xaa_1—Xaa_2—Xaa_3—Xaa_4—Z)^{A_n}$$

wherein

Xaa_1 and Xaa_3 are amino acid structures;

Xaa_2 is a valine structure;

Xaa_4 is a phenylalanine structure;

Y, which may or may not be present, is a peptidic structure having the formula $(Xaa)_a$, wherein Xaa is any amino acid structure and a is an integer from 1 to 15;

Z, which may or may not be present, is a peptidic structure having the formula $(Xaa)_b$, wherein Xaa is any amino acid structure and b is an integer from 1 to 15; and

A is a modifying group attached directly or indirectly to the compound and n is an integer;

Xaa_1, Xaa_3, Y, Z, A and n being selected such that the compound modulates the aggregation or inhibits the neurotoxicity of natural β-amyloid peptides when contacted with the natural β-amyloid peptides.

Preferably, a modulator compound of the above formula inhibits aggregation of natural β-amyloid peptides when contacted with the natural β-amyloid peptides and/or inhibits Aβ neurotoxicity. Alternatively, the modulator compound can promote aggregation of natural β-amyloid peptides when contacted with the natural β-amyloid peptides. The type and number of modifying groups ("A") coupled to the modulator are selected such that the compound alters (and preferably inhibits) aggregation of natural β-amyloid peptides when contacted with the natural β-amyloid peptides. A single modifying group can be coupled to the modulator (i.e., n=1 in the above formula) or, alternatively, multiple modifying groups can be coupled to the modulator. In various embodiments, n is an integer between 1 and 60, between 1 and 30, between 1 and 10, between 1 and 5 or between 1 and 3. Suitable types of modifying groups are described further in subsection II below.

As demonstrated in Example 9, amino acid positions 18 (Val_{18}) and 20 (Phe_{20}) of $A\beta_{17-20}$ (corresponding to Xaa_2 and Xaa_4) are particularly important within the core domain for inhibitory activity of the modulator compound. Accordingly, these positions are conserved within the core domain in the formula shown above. The terms "valine structure" and "phenylalanine structure" as used in the above formula are intended to include the natural amino acids, as well as non-naturally-occurring analogues, derivatives and mimetics of valine and phenylalanine, respectively, (including D-amino acids) which maintain the functional activity of the compound. Moreover, although Val_{18} and Phe_{20} have an important functional role, it is possible that Xaa_2 and/or Xaa_4 can be substituted with other naturally-occurring amino acids that are structurally related to valine or phenylalanine, respectively, while still maintaining the activity of the compound. Thus, the terms "valine structure" is intended to include conservative amino acid substitutions that retain the activity of valine at Xaa_2, and the term "phenylalanine structure" is intended to include conservative amino acid substitutions that retain the activity of phenylalanine at Xaa_4. However, the term "valine structure" is not intended to include threonine.

In contrast to positions 18 and 20 of $A\beta_{17-20}$, a Phe to Ala substitution at position 19 (corresponding to Xaa_3) did not

abolish the activity of the modulator, indicating position 19 may be more amenable to amino acid substitution. In various embodiments of the above formula, positions Xaa_1 and Xaa_3 are any amino acid structure. The term "amino acid structure" is intended to include natural and non-natural amino acids as well as analogues, derivatives and mimetics thereof, including D-amino acids. In a preferred embodiment of the above formula, Xaa_1 is a leucine structure and Xaa_3 is a phenylalanine structure (i.e., modeled after Leu_{17} and Phe_{19}, respectively, in the natural Aβ peptide sequence). The term "leucine structure" is used in the same manner as valine structure and phenylalanine structure described above. Alternatively, an another embodiment, Xaa_3 is an alanine structure.

The four amino acid structure ACD of the modulator of the above formula can be flanked at the amino-terminal side, carboxy-terminal side, or both, by peptidic structures derived either from the natural Aβ peptide sequence or from non-Aβ sequences. The term "peptidic structure" is intended to include peptide analogues, derivatives and mimetics thereof, as described above. The peptidic structure is composed of one or more linked amino acid structures, the type and number of which in the above formula are variable. For example, in one embodiment, no additional amino acid structures flank the Xaa_1-Xaa_2-Xaa_3-Xaa_4 core sequence (i.e., Y and Z are absent in the above formula). In another embodiment, one or more additional amino acid structures flank only the amino-terminus of the core sequences (i.e., Y is present but Z is absent in the above formula). In yet another embodiment, one or more additional amino acid structures flank only the carboxy-terminus of the core sequences (i.e., Z is present but Y is absent in the above formula). The length of flanking Z or Y sequences also is variable. For example, in one embodiment, a and b are integers from 1 to 15. More preferably, a and b are integers between 1 and 10. Even more preferably, a and b are integers between 1 and 5. Most preferably, a and b are integers between 1 and 3.

One form of the β-amyloid modulator compound comprising an Aβ aggregation core domain modeled after $Aβ_{17-20}$ coupled directly or indirectly to at least one modifying group has the formula:

A-(Y)-Xaa_1-Xaa_2-Xaa_3-Xaa_4-(Z)-B

wherein

Xaa_1 and Xaa_3 are amino acids or amino acid mimetics;

Xaa_2 is valine or a valine mimetic

Xaa_4 is phenylalanine or a phenylalanine mimetic;

Y, which may or may not be present, is a peptide or peptidomimetic having the formula $(Xaa)_a$, wherein Xaa is any amino acid or amino acid mimetic and a is an integer from 1 to 15;

Z, which may or may not be present, is a peptide or peptidomimetic having the formula $(Xaa)_b$, wherein Xaa is any amino acid or amino acid mimetic and b is an integer from 1 to 15; and

A and B, at least one of which is present, are modifying groups attached directly or indirectly to the amino terminus and carboxy terminus, respectively, of the compound;

Xaa_1, Xaa_3, Y, Z, A and B being selected such that the compound modulates the aggregation or inhibits the neurotoxicity of natural β-amyloid peptides when contacted with the natural β-amyloid peptides.

In this embodiment, the modulator compound is specifically modified at either its amino-terminus, its carboxy-terminus, or both. The terminology used in this formula is the same as described above. Suitable modifying groups are described in subsection II below. In one embodiment, the compound is modified only at its amino terminus (i.e., B is absent and the compound comprises the formula: A-(Y)-Xaa_1-Xaa_2-Xaa_3-Xaa_4-(Z)). In another embodiment, the compound is modified only at its carboxy-terminus (i.e., A is absent and the compound comprises the formula: (Y)-Xaa_1-Xaa_2-Xaa_3-Xaa_4-(Z)-B). In yet another embodiment, the compound is modified at both its amino- and carboxy termini (i.e., the compound comprises the formula: A-(Y)-Xaa_1-Xaa_2-Xaa_3-Xaa_4-(Z)-B and both A and B are present). As described above, the type and number of amino acid structures which flank the Xaa_1-Xaa_2-Xaa_3-Xaa_4 core sequences in the above formula is variable. For example, in one embodiment, a and b are integers from 1 to 15. More preferably, a and b are integers between 1 and 10. Even more preferably, a and b are integers between 1 and 5. Most preferably, a and b are integers between 1 and 3.

As demonstrated in Examples 7, 8 and 9, preferred Aβ modulator compounds of the invention comprise modified forms of $Aβ_{14-21}$ (His-Gln-Lys-Leu-Val-Phe-Phe-Ala; SEQ ID NO:5), or amino-terminal or carboxy-terminal deletions thereof, with a preferred "minimal core region" comprising $Aβ_{17-20}$. Accordingly, in specific embodiments, the invention provides compounds comprising the formula:

A-Xaa_1-Xaa_2-Xaa_3-Xaa_4-Xaa_5-Xaa_6-Xaa_7-Xaa_8-B

wherein

Xaa1 is a histidine structure;

Xaa2 is a glutamine structure;

Xaa3 is a lysine structure;

Xaa4 is a leucine structure;

Xaa5 is a valine structure;

Xaa6 is a phenylalanine structure;

Xaa7 is a phenylalanine structure;

Xaa8 is an alanine structure;

A and B are modifying groups attached directly or indirectly to the amino terminus and carboxy terminus, respectively, of the compound;

and wherein Xaa_1-Xaa_2-Xaa_3, Xaa_1-Xaa_2 or Xaa_1 may or may not be present;

Xaa_8 may or may not be present; and

at least one of A and B is present.

In one specific embodiment, the compound comprises the formula: A-Xaa_4-Xaa_5-Xaa_6-Xaa_7-B (e.g, a modified form of $Aβ_{17-20}$, comprising an amino acid sequence Leu-Val-Phe-Phe; SEQ ID NO:12).

In another specific embodiment, the compound comprises the formula: A-Xaa_4-Xaa_5-Xaa_6-Xaa_7-Xaa_8-B (e.g, a modified form of $Aβ_{17-21}$, comprising an amino acid sequence Leu-Val-Phe-Phe-Ala; SEQ ID NO:11).

In another specific embodiment, the compound comprises the formula: A-Xaa_3-Xaa_4-Xaa_5-Xaa_6-Xaa_7-B (e.g., a modified form of $Aβ_{16-20}$, comprising an amino acid sequence Lys-Leu-Val-Phe-Phe; SEQ ID NO:10).

In another specific embodiment, the compound comprises the formula: A-Xaa_3-Xaa_4-Xaa_5-Xaa_6-Xaa_7-Xaa_8-B (e.g., a modified form of $Aβ_{16-21}$, comprising an amino acid sequence Lys-Leu-Val-Phe-Phe-Ala; SEQ ID NO:9).

In another specific embodiment, the compound comprises the formula: A-Xaa_2-Xaa_3-Xaa_4-Xaa_5-Xaa_6-Xaa_7-B (e.g., a modified form of $Aβ_{15-20}$, comprising an amino acid sequence Gln-Lys-Leu-Val-Phe-Phe; SEQ ID NO:8).

In another specific embodiment, the compound comprises the formula: A-Xaa_2-Xaa_3-Xaa_4-Xaa_5-Xaa_6-Xaa_7-Xaa_8-B

(e.g., a modified form of $A\beta_{15-21}$, comprising an acid sequence Gln-Lys-Leu-Val-Phe-Phe-Ala; SEQ ID NO:7).

In another specific embodiment, the compound comprises the formula: A-Xaa_1-Xaa_2-Xaa_3-Xaa_4-Xaa_5-Xaa_6-Xaa_7-B (e.g., a modified form of $A\beta_{14-20}$, comprising an amino acid sequence His-Gln-Lys-Leu-Val-Phe-Phe; SEQ ID NO:6).

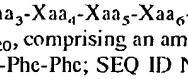
In another specific embodiment, the compound comprises the formula: A-Xaa_1-Xaa_2-Xaa_3-Xaa_4-Xaa_5-Xaa_6-Xaa_7-Xaa_8-B (e.g., a modified form of $A\beta_{14-21}$, comprising an amino acid sequence His-Gln-Lys-Leu-Val-Phe-Phe-Ala; SEQ ID NO:5).

In preferred embodiments of the aforementioned specific embodiments, A or B is a cholanoyl structure or a biotin-containing structure (described further in subsection II below).

In further experiments to delineate subregions of Aβ upon which an Aβ aggregation core domain can be modeled (the results of which are described in Example 11), it was demonstrated that a modulator compound having inhibitory activity can comprise as few as three Aβ amino acids residues (e.g., Val-Phe-Phe, which corresponds to $A\beta_{8-20}$ or Phe-Phe-Ala, which corresponds to $A\beta_{19-21}$). The results also demonstrated that a modulator compound having a modulating group at its carboxy-terminus is effective at inhibiting Aβ aggregation. Still further, the results demonstrated that the cholyl group, as a modulating group, can be manipulated while maintaining the inhibitory activity of the compounds and that an iodotyrosyl can be substituted for phenylalanine (e.g., at position 19 or 20 of the Aβ sequence) while maintaining the ability of the compound to inhibit Aβ aggregation.

Still further, the results demonstrated that compounds with inhibitory activity can be created using amino acids residues that are derived from the Aβ sequence in the region of about positions 17–21 but wherein the amino acid sequence is rearranged or has a substitution with a non-Aβ-derived amino acid. Examples of such compounds include PPI-426, in which the sequence of $A\beta_{17-21}$ (LVFFA SEQ ID NO:11) has been rearranged (FFVLA SEQ ID NO:21), PPI-372, in which the sequence of $A\beta_{16-20}$ (KLVFF SEQ ID NO:10) has been rearranged (FKFVL SEQ ID NO:29), and PPI-388, -389 and -390, in which the sequence of $A\beta_{17-21}$ (LVFFA SEQ ID NO:11) has been substituted at position 17, 18 or 19, respectively, with an alanine residue (AVFFA SEQ ID NO:25 for PPI-388, LAFFA SEQ ID NO:13 for PPI-389 and LVAFA SEQ ID NO:33 for PPI-390). The inhibitory activity of these compounds indicate that the presence in the compound of an amino acid sequence directly corresponding to a portion of Aβ is not essential for inhibitory activity, but rather suggests that maintenance of the hydrophobic nature of this core region, by inclusion of amino acid residues such as phenylalanine, valine, leucine, regardless of their precise order, can be sufficient for inhibition of Aβ aggregation. Accordingly, an Aβ aggregation core domain can be designed based on the direct Aβ amino acid sequence or can be designed based on a rearranged Aβ sequence which maintains the hydrophobicity of the Aβ subregion, e.g., the region around positions 17–20. This region of Aβ contains the amino acid residues Leu, Val and Phe. Accordingly, preferred Aβ aggregation core domains are composed of at least three amino acid structures (as that term is defined hereinbefore, including amino acid derivatives, analogues and mimetics), wherein at least two of the amino acid structures are, independently, either a leucine structure, a valine structure or a phenylalanine structure (as those terms are defined hereinbefore, including derivatives, analogues and mimetics).

Thus, in another embodiment, the invention provides a β-amyloid modulator compound comprising a formula:



wherein Xaa_1, Xaa_2 and Xaa_3 are each amino acid structures and at least two of Xaa_1, Xaa_2 and Xaa_3 are, independently, selected from the group consisting of a leucine structure, a phenylalanine structure and a valine structure;

Y, which may or may not be present, is a peptidic structure having the formula $(Xaa)_a$, wherein Xaa is any amino acid structure and a is an integer from 1 to 15;

Z, which may or may not be present, is a peptidic structure having the formula $(Xaa)_b$, wherein Xaa is any amino acid structure and b is an integer from 1 to 15; and

A is a modifying group attached directly or indirectly to the compound and n is an integer;

Xaa_1, Xaa_2, Xaa_3, Y, Z, A and n being selected such that the compound modulates the aggregation or inhibits the neurotoxicity of natural β-amyloid peptides when contacted with the natural β-amyloid peptides.

Preferably, the compound inhibits aggregation of natural β-amyloid peptides when contacted with the natural β-amyloid peptides. In preferred embodiments, Xaa_1 and Xaa_2 are each phenylalanine structures or Xaa_2 and Xaa_3 are each phenylalanine structures. "n" can be, for example, an integer between 1 and 5, whereas "a" and "b" can be, for example, integers between 1 and 5. The modifying group "A" preferably comprises a cyclic, heterocyclic or polycyclic group. More preferably, A contains a cis-decalin group, such as cholanoyl structure or a cholyl group. In other embodiments, A can comprise a biotin-containing group, a diethylene-triaminepentaacetyl group, a (−)-menthoxyacetyl group, a fluorescein-containing group or an N-acetylneuraminyl group. In yet other embodiments, the compound may promotes aggregation of natural β-amyloid peptides when contacted with the natural β-amyloid peptides, may be further modified to alter a pharmacokinetic property of the compound or may be further modified to label the compound with a detectable substance.

In another embodiment, the invention provides a β-amyloid modulator compound comprising a formula:

A-(Y)-Xaa_1-Xaa_2-Xaa_3-(Z)-B

wherein Xaa_1, Xaa_2 and Xaa_3 are each amino acid structures and at least two of Xaa_1, Xaa_2 and Xaa_3 are, independently, selected from the group consisting of a leucine structure, a phenylalanine structure and a valine structure;

Y, which may or may not be present, is a peptidic structure having the formula $(Xaa)_a$, wherein Xaa is any amino acid structure and a is an integer from 1 to 15;

Z, which may or may not be present, is a peptidic structure having the formula $(Xaa)_b$, wherein Xaa is any amino acid structure and b is an integer from 1 to 15; and

A and B, at least one of which is present, are modifying groups attached directly or indirectly to the amino terminus and carboxy terminus, respectively, of the compound;

Xaa_1, Xaa_2, Xaa_3, Y, Z, A and B being selected such that the compound modulates the aggregation or inhibits the neurotoxicity of natural β-amyloid peptides when contacted with the natural β-amyloid peptides.

Preferably, the compound inhibits aggregation of natural β-amyloid peptides when contacted with the natural β-amyloid peptides. In preferred embodiments, Xaa_1 and

Xaa_2 are each phenylalanine structures or Xaa_2 and Xaa_3 are each phenylalanine structures. In one subembodiment, the compound comprises the formula:

A-(Y)-Xaa_1-Xaa_2-Xaa_3-(Z)

In another subembodiment, the compound comprises the formula:

(Y)-Xaa_1-Xaa_2-Xaa_3-(Z)-B

"n" can be, for example, an integer between 1 and 5, whereas "a" and "b" can be, for example, integers between 1 and 5. The modifying group "A" preferably comprises a cyclic, heterocyclic or polycyclic group. More preferably, A contains a cis-decalin group, such as cholanoyl structure or a cholyl group In other embodiments, A can comprise a biotin-containing group, a diethylene-triaminepentaacetyl group, a (−)-menthoxyacetyl group, a fluorescein-containing group or an N-acetylneuraminyl group. In yet other embodiments, the compound may promote aggregation of natural β-amyloid peptides when contacted with the natural β-amyloid peptides, may be further modified to alter a pharmacokinetic property of the compound or may be further modified to label the compound with a detectable substance.

In preferred specific embodiments, the invention provides a β-amyloid modulator compound comprising a modifying group attached directly or indirectly to a peptidic structure, wherein the peptidic structure comprises amino acid structures having an amino acid sequence selected from the group consisting of His-Gln-Lys-Leu-Val-Phe-Phe-Ala (SEQ ID NO:5), His-Gln-Lys-Leu-Val-Phe-Phe (SEQ ID NO:6), Gln-Lys-Leu-Val-Phe-Phe-Ala (SEQ ID NO:7), Gln-Lys-Leu-Val-Phe-Phe (SEQ ID NO:8), Lys-Leu-Val-Phe-Phe-Ala (SEQ ID NO:9), Lys-Leu-Val-Phe-Phe (SEQ ID NO:10), Leu-Val-Phe-Phe-Ala (SEQ ID NO:11), Leu-Val-Phe-Phe (SEQ ID NO:12), Leu-Ala-Phe-Phe-Ala (SEQ ID NO:13), Val-Phe-Phe (SEQ ID NO:19), Phe-Phe-Ala (SEQ ID NO:20), Phe-Phe-Val-Leu-Ala (SEQ ID NO:21), Leu-Val-Phe-Phe-Lys (SEQ ID NO:22), Leu-Val-Iodotyrosine-Phe-Ala (SEQ ID NO:23), Val-Phe-Phe-Ala (SEQ ID NO:24), Ala-Val-Phe-Phe-Ala (SEQ ID NO:25), Leu-Val-Phe-Iodotyrosine-Ala (SEQ ID NO:26), Leu-Val-Phe-Phe-Ala-Glu (SEQ ID NO:27), Phe-Phe-Val-Leu (SEQ ID NO:28), Phe-Lys-Phe-Val-Leu (SEQ ID NO:29), Lys-Leu-Val-Ala-Phe (SEQ ID NO:30), Lys-Leu-Val-Phe-Phe-βAla (SEQ ID NO:31) and Leu-Val-Phe-Phe-DAla (SEQ ID NO:32).

These specific compounds can be further modified to alter a pharmacokinetic property of the compound and/or further modified to label the compound with a detectable substance.

The modulator compounds of the invention can be incorporated into pharmaceutical compositions (described further in subsection V below) and can be used in detection and treatment methods as described further in subsection VI below.

II. Modifying Groups

Within a modulator compound of the invention, a peptidic structure (such as an Aβ derived peptide, or an Aβ aggregation core domain, or an amino acid sequence corresponding to a rearranged Aβ aggregation core domain) is coupled directly or indirectly to at least one modifying group (abbreviated as MG). In one embodiment, a modulator compounds of the invention comprising an aggregation core domain coupled to a modifying group, the compound can be illustrated schematically as MG-ACD. The term "modifying group" is intended to include structures that are directly attached to the peptidic structure (e.g., by covalent coupling), as well as those that are indirectly attached to the peptidic structure (e.g., by a stable non-covalent association or by covalent coupling to additional amino acid residues, or mimetics, analogues or derivatives thereof, which may flank the Aβ-derived peptidic structure). For example, the modifying group can be coupled to the amino-terminus or carboxy-terminus of an Aβ-derived peptidic structure, or to a peptidic or peptidomimetic region flanking the core domain. Alternatively, the modifying group can be coupled to a side chain of at least one amino acid residue of an Aβ-derived peptidic structure, or to a peptidic or peptidomimetic region flanking the core domain (e.g., through the epsilon amino group of a lysyl residue(s), through the carboxyl group of an aspartic acid residue(s) or a glutamic acid residue(s), through a hydroxy group of a tyrosyl residue (s), a serine residue(s) or a threonine residue(s) or other suitable reactive group on an amino acid side chain). Modifying groups covalently coupled to the peptidic structure can be attached by means and using methods well known in the art for linking chemical structures, including, for example, amide, alkylamino, carbamate or urea bonds.

The term "modifying group" is intended to include groups that are not naturally coupled to natural Aβ peptides in their native form. Accordingly, the term "modifying group" is not intended to include hydrogen. The modifying group(s) is selected such that the modulator compound alters, and preferably inhibits, aggregation of natural β-amyloid peptides when contacted with the natural β-amyloid peptides or inhibits the neurotoxicity of natural β-amyloid peptides when contacted with the natural β-amyloid peptides. Although not intending to be limited by mechanism, the modifying group(s) of the modulator compounds of the invention is thought to function as a key pharmacophore which is important for conferring on the modulator the ability to disrupt Aβ polymerization.

In a preferred embodiment, the modifying group(s) comprises a cyclic, heterocyclic or polycyclic group. The term "cyclic group", as used herein, is intended to include cyclic saturated or unsaturated (i.e., aromatic) group having from about 3 to 10, preferably about 4 to 8, and more preferably about 5 to 7, carbon atoms. Exemplary cyclic groups include cyclopropyl, cyclobutyl, cyclopentyl, cyclohexyl, and cyclooctyl. Cyclic groups may be unsubstituted or substituted at one or more ring positions. Thus, a cyclic group may be substituted with, e.g., halogens, alkyls, cycloalkyls, alkenyls, alkynyls, aryls, heterocycles, hydroxyls, aminos, nitros, thiols amines, imines, amides, phosphonates, phosphines, carbonyls, carboxyls, silyls, ethers, thioethers, sulfonyls, sulfonates, selenoethers, ketones, aldehydes, esters, —CF_3, —CN, or the like.

The term "heterocyclic group" is intended to include cyclic saturated or unsaturated (i.e., aromatic) group having from about 3 to 10, preferably about 4 to 8, and more preferably about 5 to 7, carbon atoms, wherein the ring structure includes about one to four heteroatoms. Heterocyclic groups include pyrrolidine, oxolane, thiolane, imidazole, oxazole, piperidine, piperazine, morpholine. The heterocyclic ring can be substituted at one or more positions with such substituents as, for example, halogens, alkyls, cycloalkyls, alkenyls, alkynyls, aryls, other heterocycles, hydroxyl, amino, nitro, thiol, amines, imines, amides, phosphonates, phosphines, carbonyls, carboxyls, silyls, ethers, thioethers, sulfonyls, selenoethers, ketones, aldehydes, esters, —CF_3, —CN, or the like. Heterocycles may also be bridged or fused to other cyclic groups as described below.

The term "polycyclic group" as used herein is intended to refer to two or more saturated or unsaturated (i.e., aromatic)

cyclic rings in which two or more carbons are common to two adjoining rings, e.g., the rings are "fused rings". Rings that are joined through non-adjacent atoms are termed "bridged" rings. Each of the rings of the polycyclic group can be substituted with such substituents as described above, as for example, halogens, alkyls, cycloalkyls, alkenyls, alkynyls, hydroxyl, amino, nitro, thiol, amines, imines, amides, phosphonates, phosphines, carbonyls, carboxyls, silyls, ethers, thioethers, sulfonyls, selenoethers, ketones, aldehydes, esters, —CF_3, —CN, or the like.

A preferred polycyclic group is a group containing a cis-decalin structure. Although not intending to be limited by mechanism, it is thought that the "bent" conformation conferred on a modifying group by the presence of a cis-decalin structure contributes to the efficacy of the modifying group in disrupting Aβ polymerization. Accordingly, other structures which mimic the "bent" configuration of the cis-decalin structure can also be used as modifying groups. An example of a cis-decalin containing structure that can be used as a modifying group is a cholanoyl structure, such as a cholyl group. For example, a modulator compound can be modified at its amino terminus with a cholyl group by reacting the aggregation core domain with cholic acid, a bile acid, as described in Example 4 (the structure of cholic acid is illustrated in FIG. 2). Moreover, a modulator compound can be modified at its carboxy terminus with a cholyl group according to methods known in the art (see e.g., Wess, G. et al. (1993) *Tetrahedron Letters,* 34:817–822; Wess, G. et al. (1992) *Tetrahedron Letters* 33:195–198; and Kramer, W. et al. (1992) *J. Biol. Chem.* 267:18598–18604). Cholyl derivatives and analogues can also be used as modifying groups. For example, a preferred cholyl derivative is Aic (3-(O-aminoethyl-iso)-cholyl), which has a free amino group that can be used to further modify the modulator compound (e.g., a chelation group for ^{99m}Tc can be introduced through the free amino group of Aic). As used herein, the term "cholanoyl structure" is intended to include the cholyl group and derivatives and analogues thereof, in particular those which retain a four-ring cis-decalin configuration. Examples of cholanoyl structures include groups derived from other bile acids, such as deoxycholic acid, lithocholic acid, ursodeoxycholic acid, chenodeoxycholic acid and hyodeoxycholic acid, as well as other related structures such as cholanic acid, bufalin and resibufogenin (although the latter two compounds are not preferred for use as a modifying group). Another example of a cis-decalin containing compound is 5β-cholestan-3α-ol (the cis-decalin isomer of (+)-dihydrocholesterol). For further description of bile acid and steroid structure and nomenclature, see Nes, W. R. and McKean, M. L. *Biochemistry of Steroids and Other Isopentanoids,* University Park Press, Baltimore, Md., Chapter 2.

In addition to cis-decalin containing groups, other polycyclic groups may be used as modifying groups. For example, modifying groups derived from steroids or β-lactams may be suitable modifying groups. Moreover, non-limiting examples of some additional cyclic, heterocyclic or polycyclic compounds which can be used to modify an Aβ-derived peptidic structure are shown schematically in FIG. 2. In one embodiment, the modifying group is a "biotinyl structure", which includes biotinyl groups and analogues and derivatives thereof (such as a 2-iminobiotinyl group). In another embodiment, the modifying group can comprise a "fluorescein-containing group", such as a group derived from reacting an Aβ-derived peptidic structure with 5-(and 6-)-carboxyfluorescein, succinimidyl ester or fluorescein isothiocyanate. In various other embodiments, the modifying group(s) can comprise an N-acetylneuraminyl group, a trans-4-cotininecarboxyl group, a 2-imino-1-imidazolidineacetyl group, an (S)-(−)-indoline-2-carboxyl group, a (−)-menthoxyacetyl group, a 2-norbornaneacetyl group, a γ-oxo-5-acenaphthenebutyryl, a (−)-2-oxo-4-thiazolidinecarboxyl group, a tetrahydro-3-furoyl group, a 2-iminobiotinyl group, a diethylenetriaminepentaacetyl group, a 4-morpholinecarbonyl group, a 2-thiopheneacetyl group or a 2-thiophenesulfonyl group.

Preferred modifying groups include groups comprising cholyl structures, biotinyl structures, fluorescein-containing groups, a diethylene-triaminepentaacetyl group, a (−)-menthoxyacetyl group, and a N-acetylneuraminyl group. More preferred modifying groups those comprising a cholyl structure or an iminiobiotinyl group.

In addition to the cyclic, heterocyclic and polycyclic groups discussed above, other types of modifying groups can be used in a modulator of the invention. For example, small hydrophobic groups may be suitable modifying groups. An example of a suitable non-cyclic modifying group is an acetyl group.

Yet another type of modifying group is a compound that contains a non-natural amino acid that acts as a beta-turn mimetic, such as a dibenzofuran-based amino acid described in Tsang, K. Y. et al. (1994) *J. Am. Chem. Soc.* 116:3988–4005; Diaz, H. and Kelly, J. W. (1991) *Tetrahedron Letters* 41:5725–5728; and Diaz, H. et al. (1992) *J. Am. Chem. Soc.* 114:8316–8318. An example of such a modifying group is a peptide-aminoethyldibenzofuranyl-proprionic acid (Adp) group (e.g., DDIIL-Adp; SEQ ID NO:34). This type of modifying group further can comprise one or more N-methyl peptide bonds to introduce additional steric hindrance to the aggregation of natural β-AP when compounds of this type interact with natural β-AP.

III. Additional Chemical Modifications of Aβ Modulators

A β-amyloid modulator compound of the invention can be further modified to alter the specific properties of the compound while retaining the ability of the compound to alter Aβ aggregation and inhibit Aβ neurotoxicity. For example, in one embodiment, the compound is further modified to alter a pharmacokinetic property of the compound, such as in vivo stability or half-life. In another embodiment, the compound is further modified to label the compound with a detectable substance. In yet another embodiment, the compound is further modified to couple the compound to an additional therapeutic moiety. Schematically, a modulator of the invention comprising an Aβ aggregation core domain coupled directly or indirectly to at least one modifying group can be illustrated as MG-ACD, whereas this compound which has been further modified to alter the properties of the modulator can be illustrated as MG-ACD-CM, wherein CM represents an additional chemical modification.

To further chemically modify the compound, such as to alter the pharmacokinetic properties of the compound, reactive groups can be derivatized. For example, when the modifying group is attached to the amino-terminal end of the aggregation core domain, the carboxy-terminal end of the compound can be further modified. Preferred C-terminal modifications include those which reduce the ability of the compound to act as a substrate for carboxypeptidases. Examples of preferred C-terminal modifiers include an amide group, an ethylamide group and various non-natural amino acids, such as D-amino acids and β-alanine. Alternatively, when the modifying group is attached to the carboxy-terminal end of the aggregation core domain, the amino-terminal end of the compound can be further modified, for example, to reduce the ability of the compound to act as a substrate for aminopeptidases.

A modulator compound can be further modified to label the compound by reacting the compound with a detectable substance. Suitable detectable substances include various enzymes, prosthetic groups, fluorescent materials, luminescent materials and radioactive materials. Examples of suitable enzymes include horseradish peroxidase, alkaline phosphatase, β-galactosidase, or acetylcholinesterase; examples of suitable prosthetic group complexes include streptavidin/biotin and avidin/biotin; examples of suitable fluorescent materials include umbelliferone, fluorescein, fluorescein isothiocyanate, rhodamine, dichlorotriazinylamine fluorescein, dansyl chloride or phycoerythrin; an example of a luminescent material includes luminol; and examples of suitable radioactive material include ^{14}C, ^{123}I, ^{124}I, ^{125}I, ^{131}I, ^{99m}Tc, ^{35}S or ^{3}H. In a preferred embodiment, a modulator compound is radioactively labeled with ^{14}C, either by incorporation of ^{14}C into the modifying group or one or more amino acid structures in the modulator compound. Labeled modulator compounds can be used to assess the in vivo pharmacokinetics of the compounds, as well as to detect Aβ aggregation, for example for diagnostic purposes. Aβ aggregation can be detected using a labeled modulator compound either in vivo or in an in vitro sample derived from a subject.

Preferably, for use as an in vivo diagnostic agent, a modulator compound of the invention is labeled with radioactive technetium or iodine. Accordingly, in one embodiment, the invention provides a modulator compound labeled with technetium, preferably ^{99m}Tc. Methods for labeling peptide compounds with technetium are known in the art (see e.g., U.S. Pat. Nos. 5,443,815, 5,225,180 and 5,405,597, all by Dean et al.; Stepniak-Biniakiewicz, D., et al. (1992) *J. Med. Chem.* 35:274–279; Fritzberg, A. R., et al. (1988) *Proc. Natl. Acad. Sci. USA* 85:4025–4029; Baidoo, K. E., et al. (1990) *Cancer Res. Suppl.* 50:799s–803s; and Regan, L. and Smith, C. K. (1995) *Science* 270:980–982). A modifying group can be chosen that provides a site at which a chelation group for ^{99m}Tc can be introduced, such as the Aic derivative of cholic acid, which has a free amino group (see Example 11). In another embodiment, the invention provides a modulator compound labeled with radioactive iodine. For example, a phenylalanine residue within the Aβ sequence (such as Phe_{19} or Phe_{20}) can be substituted with radioactive iodotyrosyl (see Example 11). Any of the various isotopes of radioactive iodine can be incorporated to create a diagnostic agent. Preferably, ^{123}I (half-life=13.2 hours) is used for whole body scintigraphy, ^{124}I (half life=4 days) is used for positron emission tomography (PET), ^{125}I (half life=60 days) is used for metabolic turnover studies and 131I (half life=8 days) is used for whole body counting and delayed low resolution imaging studies.

Furthermore, an additional modification of a modulator compound of the invention can serve to confer an additional therapeutic property on the compound. That is, the additional chemical modification can comprise an additional functional moiety. For example, a functional moiety which serves to break down or dissolve amyloid plaques can be coupled to the modulator compound. In this form, the MG-ACD portion of the modulator serves to target the compound to Aβ peptides and disrupt the polymerization of the Aβ peptides, whereas the additional functional moiety serves to break down or dissolve amyloid plaques after the compound has been targeted to these sites.

In an alternative chemical modification, a β-amyloid compound of the invention is prepared in a "prodrug" form, wherein the compound itself does not modulate Aβ aggregation, but rather is capable of being transformed, upon metabolism in vivo, into a β-amyloid modulator compound as defined herein. For example, in this type of compound, the modulating group can be present in a prodrug form that is capable of being converted upon metabolism into the form of an active modulating group. Such a prodrug form of a modifying group is referred to herein as a "secondary modifying group." A variety of strategies are known in the art for preparing peptide prodrugs that limit metabolism in order to optimize delivery of the active form of the peptide-based drug (see e.g., Moss, J. (1995) in *Peptide-Based Drug Design: Controlling Transport and Metabolism*, Taylor, M. D. and Amidon, G. L. (eds), Chapter 18. Additionally strategies have been specifically tailored to achieving CNS delivery based on "sequential metabolism" (see e.g., Bodor, N., et al. (1992) *Science* 257:1698–1700; Prokai, L., et al. (1994) *J. Am. Chem. Soc.* 116:2643–2644; Bodor, N. and Prokai, L. (1995) in *Peptide-Based Drug Design: Controlling Transport and Metabolism*, Taylor, M. D. and Amidon, G. L. (eds), Chapter 14. In one embodiment of a prodrug form of a modulator of the invention, the modifying group comprises an alkyl ester to facilitate blood-brain barrier permeability.

Modulator compounds of the invention can be prepared by standard techniques known in the art. The peptide component of a modulator composed, at least in part, of a peptide, can be synthesized using standard techniques such as those described in Bodansky, M. *Principles of Peptide Synthesis*, Springer Verlag, Berlin (1993) and Grant, G. A. (ed.). *Synthetic Peptides: A User's Guide*, W. H. Freeman and Company, New York (1992). Automated peptide synthesizers are commercially available (e.g., Advanced ChemTech Model 396; Milligen/Biosearch 9600). Additionally, one or more modulating groups can be attached to the Aβ-derived peptidic component (e.g., an Aβ aggregation core domain) by standard methods, for example using methods for reaction through an amino group (e.g., the alpha-amino group at the amino-terminus of a peptide), a carboxyl group (e.g., at the carboxy terminus of a peptide), a hydroxyl group (e.g., on a tyrosine, serine or threonine residue) or other suitable reactive group on an amino acid side chain (see e.g., Greene, T. W. and Wuts, P. G. M. *Protective Groups in Organic Synthesis*, John Wiley and Sons, Inc., New York (1991)). Exemplary syntheses of preferred β amyloid modulators is described further in Examples 1, 4 and 11.

IV. Screening Assays

Another aspect of the invention pertains to a method for selecting a modulator of β-amyloid aggregation. In the method, a test compound is contacted with natural β amyloid peptides, the aggregation of the natural β-AP is measured and a modulator is selected based on the ability of the test compound to alter the aggregation of the natural β-AP (e.g., inhibit or promote aggregation). In a preferred embodiment, the test compound is contacted with a molar excess amount of the natural β-AP. The amount and/or rate of natural β-AP aggregation in the presence of the test compound can be determined by a suitable assay indicative of β-AP aggregation, as described herein (see e.g., Examples 2, 5 and 6).

In a preferred assay, the natural β-AP is dissolved in solution in the presence of the test compound and aggregation of the natural β-AP is assessed in a nucleation assay (see Example 6) by assessing the turbidity of the solution over time, as measured by the apparent absorbance of the solution at 405 nm (described further in Example 6; see also Jarrett et al. (1993) *Biochemistry* 32:4693–4697). In the absence of a β-amyloid modulator, the $A_{405\ nm}$ of the solution typically

stays relatively constant during a lag time in which the β-AP remains in solution, but then the $A_{405\ nm}$ of the solution rapidly increases as the β-AP aggregates and comes out of solution, ultimately reaching a plateau level (i.e., the $A_{405\ nm}$ of the solution exhibits sigmoidal kinetics over time). In contrast, in the presence of a test compound that inhibits β-AP aggregation, the $A_{405\ nm}$, of the solution is reduced compared to when the modulator is absent. Thus, in the presence of the inhibitory modulator, the solution may exhibit an increased lag time, a decreased slope of aggregation and/or a lower plateau level compared to when the modulator is absent. This method for selecting a modulator of β-amyloid polymerization can similarly be used to select modulators that promote β-AP aggregation. Thus, in the presence of a modulator that promotes β-AP aggregation, the $A_{405\ nm}$ of the solution is increased compared to when the modulator is absent (e.g., the solution may exhibit an decreased lag time, increase slope of aggregation and/or a higher plateau level compared to when the modulator is absent).

Another assay suitable for use in the screening method of the invention, a seeded extension assay, is also described further in Example 6. In this assay, β-AP monomer and an aggregated β-AP "seed" are combined, in the presence and absence of a test compound, and the amount of β-fibril formation is assayed based on enhanced emission of the dye Thioflavine T when contacted with β-AP fibrils. Moreover, β-AP aggregation can be assessed by electron microscopy (EM) of the β-AP preparation in the presence or absence of the modulator. For example, β amyloid fibril formation, which is detectable by EM, is reduced in the presence of a modulator that inhibits β-AP aggregation (i.e., there is a reduced amount or number of β-fibrils in the presence of the modulator), whereas β fibril formation is increased in the presence of a modulator that promotes β-AP aggregation (i.e., there is an increased amount or number of β-fibrils in the presence of the modulator).

An even more preferred assay for use in the screening method of the invention to select suitable modulators is the neurotoxicity assay described in Examples 3 and 10. Compounds are selected which inhibit the formation of neurotoxic Aβ aggregates and/or which inhibit the neurotoxicity of preformed Aβ fibrils. This neurotoxicity assay is considered to be predictive of neurotoxicity in vivo. Accordingly, inhibitory activity of a modulator compound in the in vitro neurotoxicity assay is predictive of similar inhibitory activity of the compound for neurotoxicity in vivo.

V. Pharmaceutical Compositions

Another aspect of the invention pertains to pharmaceutical compositions of the β-amyloid modulator compounds of the invention. In one embodiment, the composition includes a β amyloid modulator compound in a therapeutically or prophylactically effective amount sufficient to alter, and preferably inhibit, aggregation of natural β-amyloid peptides, and a pharmaceutically acceptable carrier. In another embodiment, the composition includes a β amyloid modulator compound in a therapeutically or prophylactically effective amount sufficient to inhibit the neurotoxicity of natural β-amyloid peptides, and a pharmaceutically acceptable carrier. A "therapeutically effective amount" refers to an amount effective, at dosages and for periods of time necessary, to achieve the desired therapeutic result, such as reduction or reversal or β-amyloid deposition and/or reduction or reversal of Aβ neurotoxicity. A therapeutically effective amount of modulator may vary according to factors such as the disease state, age, sex, and weight of the individual, and the ability of the modulator to elicit a desired response in the individual. Dosage regimens may be adjusted to provide the optimum therapeutic response. A therapeutically effective amount is also one in which any toxic or detrimental effects of the modulator are outweighed by the therapeutically beneficial effects. The potential neurotoxicity of the modulators of the invention can be assayed using the cell-based assay described in Examples 3 and 10 and a therapeutically effective modulator can be selected which does not exhibit significant neurotoxicity. In a preferred embodiment, a therapeutically effective amount of a modulator is sufficient to alter, and preferably inhibit, aggregation of a molar excess amount of natural β-amyloid peptides. A "prophylactically effective amount" refers to an amount effective, at dosages and for periods of time necessary, to achieve the desired prophylactic result, such as preventing or inhibiting the rate of β-amyloid deposition and/or Aβ neurotoxicity in a subject predisposed to β-amyloid deposition. A prophylactically effective amount can be determined as described above for the therapeutically effective amount. Typically, since a prophylactic dose is used in subjects prior to or at an earlier stage of disease, the prophylactically effective amount will be less than the therapeutically effective amount.

One factor that may be considered when determining a therapeutically or prophylactically effective amount of a β amyloid modulator is the concentration of natural β-AP in a biological compartment of a subject, such as in the cerebrospinal fluid (CSF) of the subject. The concentration of natural β-AP in the CSF has been estimated at 3 nM (Schwartzman, (1994) *Proc. Natl. Acad. Sci. USA* 91:8368–8372). A non-limiting range for a therapeutically or prophylactically effective amounts of a β amyloid modulator is 0.01 nM–10 μM. It is to be noted that dosage values may vary with the severity of the condition to be alleviated. It is to be further understood that for any particular subject, specific dosage regimens should be adjusted over time according to the individual need and the professional judgment of the person administering or supervising the administration of the compositions, and that dosage ranges set forth herein are exemplary only and are not intended to limit the scope or practice of the claimed composition.

The amount of active compound in the composition may vary according to factors such as the disease state, age, sex, and weight of the individual, each of which may affect the amount of natural β-AP in the individual. Dosage regimens may be adjusted to provide the optimum therapeutic response. For example, a single bolus may be administered, several divided doses may be administered over time or the dose may be proportionally reduced or increased as indicated by the exigencies of the therapeutic situation. It is especially advantageous to formulate parenteral compositions in dosage unit form for ease of administration and uniformity of dosage. Dosage unit form as used herein refers to physically discrete units suited as unitary dosages for the mammalian subjects to be treated; each unit containing a predetermined quantity of active compound calculated to produce the desired therapeutic effect in association with the required pharmaceutical carrier. The specification for the dosage unit forms of the invention are dictated by and directly dependent on (a) the unique characteristics of the active compound and the particular therapeutic effect to be achieved, and (b) the limitations inherent in the art of compounding such an active compound for the treatment of sensitivity in individuals.

As used herein "pharmaceutically acceptable carrier" includes any and all solvents, dispersion media, coatings, antibacterial and antifungal agents, isotonic and absorption

delaying agents, and the like that are physiologically compatible. In one embodiment, the carrier is suitable for parenteral administration. Preferably, the carrier is suitable for administration into the central nervous system (e.g., intraspinally or intracerebrally). Alternatively, the carrier can be suitable for intravenous, intraperitoneal or intramuscular administration. In another embodiment, the carrier is suitable for oral administration. Pharmaceutically acceptable carriers include sterile aqueous solutions or dispersions and sterile powders for the extemporaneous preparation of sterile injectable solutions or dispersion. The use of such media and agents for pharmaceutically active substances is well known in the art. Except insofar as any conventional media or agent is incompatible with the active compound, use thereof in the pharmaceutical compositions of the invention is contemplated. Supplementary active compounds can also be incorporated into the compositions.

Therapeutic compositions typically must be sterile and stable under the conditions of manufacture and storage. The composition can be formulated as a solution, microemulsion, liposome, or other ordered structure suitable to high drug concentration. The carrier can be a solvent or dispersion medium containing, for example, water, ethanol, polyol (for example, glycerol, propylene glycol, and liquid polyethylene glycol, and the like), and suitable mixtures thereof. The proper fluidity can be maintained, for example, by the use of a coating such as lecithin, by the maintenance of the required particle size in the case of dispersion and by the use of surfactants. In many cases, it will be preferable to include isotonic agents, for example, sugars, polyalcohols such as mannitol, sorbitol, or sodium chloride in the composition. Prolonged absorption of the injectable compositions can be brought about by including in the composition an agent which delays absorption, for example, monostearate salts and gelatin. Moreover, the modulators can be administered in a time release formulation, for example in a composition which includes a slow release polymer. The active compounds can be prepared with carriers that will protect the compound against rapid release, such as a controlled release formulation, including implants and microencapsulated delivery systems. Biodegradable, biocompatible polymers can be used, such as ethylene vinyl acetate, polyanhydrides, polyglycolic acid, collagen, polyorthoesters, polylactic acid and polylactic, polyglycolic copolymers (PLG). Many methods for the preparation of such formulations are patented or generally known to those skilled in the art.

Sterile injectable solutions can be prepared by incorporating the active compound (e.g., β-amyloid modulator) in the required amount in an appropriate solvent with one or a combination of ingredients enumerated above, as required, followed by filtered sterilization. Generally, dispersions are prepared by incorporating the active compound into a sterile vehicle which contains a basic dispersion medium and the required other ingredients from those enumerated above. In the case of sterile powders for the preparation of sterile injectable solutions, the preferred methods of preparation are vacuum drying and freeze-drying which yields a powder of the active ingredient plus any additional desired ingredient from a previously sterile-filtered solution thereof.

A modulator compound of the invention can be formulated with one or more additional compounds that enhance the solubility of the modulator compound. Preferred compounds to be added to formulations to enhance the solubility of the modulators are cyclodextrin derivatives, preferably hydroxypropyl-γ-cyclodextrin. Drug delivery vehicles containing a cyclodextrin derivative for delivery of peptides to the central nervous system are described in Bodor, N., et al. (1992) *Science* 257:1698–1700. For the β-amyloid modulators described herein, inclusion in the formulation of hydroxypropyl-γ-cyclodextrin at a concentration 50–200 mM increases the aqueous solubility of the compounds. In addition to increased solubility, inclusion of a cyclodextrin derivative in the formulation may have other beneficial effects, since β-cyclodextrin itself has been reported to interact with the Aβ peptide and inhibit fibril formation in vitro (Camilleri, P., et al. (1994) *FEBS Letters* 341:256–258. Accordingly, use of a modulator compound of the invention in combination with a cyclodextrin derivative may result in greater inhibition of Aβ aggregation than use of the modulator alone. Chemical modifications of cyclodextrins are known in the art (Hanessian, S., et al. (1995) J. Org. Chem. 60:4786–4797). In addition to use as an additive in a pharmaceutical composition containing a modulator of the invention, cyclodextrin derivatives may also be useful as modifying groups and, accordingly, may also be covalently coupled to an Aβ peptide compound to form a modulator compound of the invention.

In another embodiment, a pharmaceutical composition comprising a modulator of the invention is formulated such that the modulator is transported across the blood-brain barrier (BBB). Various strategies known in the art for increasing transport across the BBB can be adapted to the modulators of the invention to thereby enhance transport of the modulators across the BBB (for reviews of such strategies, see e.g., Pardridge, W. M. (1994) *Trends in Biotechnol.* 12:239–245; Van Bree, J. B. et al. (1993) *Pharm. World Sci.* 15:2–9; and Pardridge, W. M. et al. (1992) *Pharmacol. Toxicol.* 71:3–10). In one approach, the modulator is chemically modified to form a prodrug with enhanced transmembrane transport. Suitable chemical modifications include covalent linking of a fatty acid to the modulator through an amide or ester linkage (see e.g., U.S. Pat. No. 4,933,324 and PCT Publication WO 89/07938, both by Shashoua; U.S. Pat. No. 5,284,876 by Hesse et al.; Toth, I. et al. (1 994) *J. Drug Target.* 2:217–239; and Shashoua, V. E. et al. (1984) *J. Med. Chem.* 27:659–664) and glycating the modulator (see e.g., U.S. Pat. No. 5,260,308 by Poduslo et al.). Also, N-acylamino acid derivatives may be used in a modulator to form a "lipidic" prodrug (see e.g., U.S. Pat. No. 5,112,863 by Hashimoto et al.).

In another approach for enhancing transport across the BBB, a peptidic or peptidomimetic modulator is conjugated to a second peptide or protein, thereby forming a chimeric protein, wherein the second peptide or protein undergoes absorptive-mediated or receptor-mediated transcytosis through the BBB. Accordingly, by coupling the modulator to this second peptide or protein, the chimeric protein is transported across the BBB. The second peptide or protein can be a ligand for a brain capillary endothelial cell receptor ligand. For example, a preferred ligand is a monoclonal antibody that specifically binds to the transferrin receptor on brain capillary endothelial cells (see e.g., U.S. Pat. Nos. 5,182,107 and 5,154,924 and PCT Publications WO 93/10819 and WO 95/02421, all by Friden et al.). Other suitable peptides or proteins that can mediate transport across the BBB include histones (see e.g., U.S. Pat. No. 4,902,505 by Pardridge and Schimmel) and ligands such as biotin, folate, niacin, pantothenic acid, riboflavin, thiamin, pyridoxal and ascorbic acid (see e.g., U.S. Pat. Nos. 5,416, 016 and 5,108,921, both by Heinstein). Additionally, the glucose transporter GLUT-1 has been reported to transport glycopeptides (L-serinyl-β-D-glucoside analogues of [Met5]enkephalin) across the BBB (Polt, R. et al. (1994)

Proc. Natl. Acad. Sci. USA 91:7114–1778). Accordingly, a modulator compound can be coupled to such a glycopeptide to target the modulator to the GLUT-1 glucose transporter. For example, a modulator compound which is modified at its amino terminus with the modifying group Aic (3-(O-aminoethyl-iso)-cholyl, a derivative of cholic acid having a free amino group) can be coupled to a glycopeptide through the amino group of Aic by standard methods. Chimeric proteins can be formed by recombinant DNA methods (e.g., by formation of a chimeric gene encoding a fusion protein) or by chemical crosslinking of the modulator to the second peptide or protein to form a chimeric protein. Numerous chemical crosslinking agents are known in the (e.g., commercially available from Pierce, Rockford Ill.). A crosslinking agent can be chosen which allows for high yield coupling of the modulator to the second peptide or protein and for subsequent cleavage of the linker to release bioactive modulator. For example, a biotin-avidin-based linker system may be used.

In yet another approach for enhancing transport across the BBB, the modulator is encapsulated in a carrier vector which mediates transport across the BBB. For example, the modulator can be encapsulated in a liposome, such as a positively charged unilamellar liposome (see e.g., PCT Publications WO 88/07851 and WO 88/07852, both by Faden) or in polymeric microspheres (see e.g., U.S. Pat. No. 5,413,797 by Khan et al., U.S. Pat. No. 5,271,961 by Mathiowitz et al. and U.S. Pat. No. 5,019,400 by Gombotz et al.). Moreover, the carrier vector can be modified to target it for transport across the BBB. For example, the carrier vector (e.g., liposome) can be covalently modified with a molecule which is actively transported across the BBB or with a ligand for brain endothelial cell receptors, such as a monoclonal antibody that specifically binds to transferrin receptors (see e.g., PCT Publications WO 91/04014 by Collins et al. and WO 94/02178 by Greig et al.).

In still another approach to enhancing transport of the modulator across the BBB, the modulator is coadministered with another agent which functions to permeabilize the BBB. Examples of such BBB "permeabilizers" include bradykinin and bradykinin agonists (see e.g., U.S. Pat. No. 5,112,596 by Malfroy-Camine) and peptidic compounds disclosed in U.S. Pat. No. 5,268,164 by Kozarich et al.

A modulator compound of the invention can be formulated into a pharmaceutical composition wherein the modulator is the only active compound or, alternatively, the pharmaceutical composition can contain additional active compounds. For example, two or more modulator compounds may be used in combination. Moreover, a modulator compound of the invention can be combined with one or more other agents that have anti-amyloidogenic properties. For example, a modulator compound can be combined with the non-specific cholinesterase inhibitor tacrine (COGNEX®, Parke-Davis).

In another embodiment, a pharmaceutical composition of the invention is provided as a packaged formulation. The packaged formulation may include a pharmaceutical composition of the invention in a container and printed instructions for administration of the composition for treating a subject having a disorder associated with β-amyloidosis, e.g. Alzheimer's disease.

VI. Methods of Using Aβ Modulators

Another aspect of the invention pertains to methods for altering the aggregation or inhibiting the neurotoxicity of natural β-amyloid peptides. In the methods of the invention, natural β amyloid peptides are contacted with a β amyloid modulator such that the aggregation of the natural β amyloid peptides is altered or the neurotoxicity of the natural β amyloid peptides is inhibited. In a preferred embodiment, the modulator inhibits aggregation of the natural β amyloid peptides. In another embodiment, the modulator promotes aggregation of the natural β amyloid peptides. Preferably, aggregation of a molar excess amount of β-AP, relative to the amount of modulator, is altered upon contact with the modulator.

In the method of the invention, natural β amyloid peptides can be contacted with a modulator either in vitro or in vivo. Thus, the term "contacted with" is intended to encompass both incubation of a modulator with a natural β-AP preparation in vitro and delivery of the modulator to a site in vivo where natural β-AP is present. Since the modulator compound interacts with natural β-AP, the modulator compounds can be used to detect natural β-AP, either in vitro or in vivo. Accordingly, one use of the modulator compounds of the invention is as diagnostic agents to detect the presence of natural β-AP, either in a biological sample or in vivo in a subject. Furthermore, detection of natural β-AP utilizing a modulator compound of the invention further can be used to diagnose amyloidosis in a subject. Additionally, since the modulator compounds of the invention disrupt β-AP aggregation and inhibit β-AP neurotoxicity, the modulator compounds also are useful in the treatment of disorders associated with β-amyloidosis, either prophylactically or therapeutically. Accordingly, another use of the modulator compounds of the invention is as therapeutic agents to alter aggregation and/or neurotoxicity of natural β-AP.

In one embodiment, a modulator compound of the invention is used in vitro, for example to detect and quantitate natural β-AP in sample (e.g., a sample of biological fluid). To aid in detection, the modulator compound can be modified with a detectable substance. The source of natural β-AP used in the method can be, for example, a sample of cerebrospinal fluid (e.g., from an AD patient, an adult susceptible to AD due to family history, or a normal adult). The natural β-AP sample is contacted with a modulator of the invention and aggregation of the β-AP is measured, such as by as assay described in Examples 2, 5 and 6. Preferably, the nucleation assay and/or seeded extension assay described in Example 6 is used. The degree of aggregation of the β-AP sample can then be compared to that of a control sample(s) of a known concentration of β-AP, similarly contacted with the modulator and the results can be used as an indication of whether a subject is susceptible to or has a disorder associated with β-amyloidosis. Moreover, β-AP can be detected by detecting a modulating group incorporated into the modulator. For example, modulators incorporating a biotin compound as described herein (e.g., an amino-terminally biotinylated β-AP peptide) can be detected using a streptavidin or avidin probe which is labeled with a detectable substance (e.g., an enzyme, such as peroxidase). Detection of natural β-AP aggregates mixed with a modulator of the invention using a probe that binds to the modulating group (e.g., biotin/streptavidin) is described further in Example 2.

In another embodiment, a modulator compound of the invention is used in vivo to detect, and, if desired, quantitate, natural β-AP deposition in a subject, for example to aid in the diagnosis of β amyloidosis in the subject. To aid in detection, the modulator compound can be modified with a detectable substance, preferably ^{99m}Tc or radioactive iodine (described further above), which can be detected in vivo in a subject. The labeled β-amyloid modulator compound is administered to the subject and, after sufficient time to allow accumulation of the modulator at sites of amyloid

deposition, the labeled modulator compound is detected by standard imaging techniques. The radioactive signal generated by the labeled compound can be directly detected (e.g., whole body counting), or alternatively, the radioactive signal can be converted into an image on an autoradiograph or on a computer screen to allow for imaging of amyloid deposits in the subject. Methods for imaging amyloidosis using radiolabeled proteins are known in the art. For example, serum amyloid β component (SAP), radiolabeled with either ^{123}I or ^{99m}Tc, has been used to image systemic amyloidosis (see e.g., Hawkins, P. N. and Pepys, M. B. (1995) *Eur. J Nucl. Med.* 22:595–599). Of the various isotypes of radioactive iodine, preferably ^{123}I (half-life=13.2 hours) is used for whole body scintigraphy, ^{124}I (half life=4 days) is used for positron emission tomography (PET), ^{125}I (half life=60 days) is used for metabolic turnover studies and ^{131}I (half life=8 days) is used for whole body counting and delayed low resolution imaging studies. Analogous to studies using radiolabeled SAP, a labeled modulator compound of the invention can be delivered to a subject by an appropriate route (e.g., intravenously, intraspinally, intracerebrally) in a single bolus, for example containing 100 μg of labeled compound carrying approximately 180 MBq of radioactivity.

The invention provides a method for detecting the presence or absence of natural β-amyloid peptides in a biological sample, comprising contacting a biological sample with a compound of the invention and detecting the compound bound to natural β-amyloid peptides to thereby detect the presence or absence of natural β-amyloid peptides in the biological sample. In one embodiment, the β-amyloid modulator compound and the biological sample are contacted in vitro. In another embodiment, the β-amyloid modulator compound is contacted with the biological sample by administering the β-amyloid modulator compound to a subject. For in vivo administration, preferably the compound is labeled with radioactive technetium or radioactive iodine.

The invention also provides a method for detecting natural β-amyloid peptides to facilitate diagnosis of a β-amyloidogenic disease, comprising contacting a biological sample with the compound of the invention and detecting the compound bound to natural β-amyloid peptides to facilitate diagnosis of a β-amyloidogenic disease. In one embodiment, the β-amyloid modulator compound and the biological sample are contacted in vitro. In another embodiment, the β-amyloid modulator compound is contacted with the biological sample by administering the β-amyloid modulator compound to a subject. For in vivo administration, preferably the compound is labeled with radioactive technetium or radioactive iodine. Preferably, use of the method facilitates diagnosis of Alzheimer's disease.

In another embodiment, the invention provides a method for altering natural β-AP aggregation or inhibiting β-AP neurotoxicity, which can be used prophylactically or therapeutically in the treatment or prevention of disorders associated with β amyloidosis, e.g., Alzheimer's Disease. As demonstrated in Example 10, modulator compounds of the invention reduce the toxicity of natural β-AP aggregates to cultured neuronal cells. Moreover, the modulators not only reduce the formation of neurotoxic aggregates but also have the ability to reduce the neurotoxicity of preformed Aβ fibrils. Accordingly, the modulator compounds of the invention can be used to inhibit or prevent the formation of neurotoxic Aβ fibrils in subjects (e.g., prophylactically in a subject predisposed to β-amyloid deposition) and can be used to reverse β-amyloidosis therapeutically in subjects already exhibiting β-amyloid deposition.

A modulator of the invention is contacted with natural β amyloid peptides present in a subject (e.g., in the cerebrospinal fluid or cerebrum of the subject) to thereby alter the aggregation of the natural β-AP and/or inhibit the neurotoxicity of the natural β-APs. A modulator compound alone can be administered to the subject, or alternatively, the modulator compound can be administered in combination with other therapeutically active agents (e.g., as discussed above in subsection IV). When combination therapy is employed, the therapeutic agents can be coadministered in a single pharmaceutical composition, coadministered in separate pharmaceutical compositions or administered sequentially.

The modulator may be administered to a subject by any suitable route effective for inhibiting natural β-AP aggregation in the subject, although in a particularly preferred embodiment, the modulator is administered parenterally, most preferably to the central nervous system of the subject. Possible routes of CNS administration include intraspinal administration and intracerebral administration (e.g., intracerebrovascular administration). Alternatively, the compound can be administered, for example, orally, intraperitoneally, intravenously or intramuscularly. For non-CNS administration routes, the compound can be administered in a formulation which allows for transport across the BBB. Certain modulators may be transported across the BBB without any additional further modification whereas others may need further modification as described above in subsection IV.

Suitable modes and devices for delivery of therapeutic compounds to the CNS of a subject are known in the art, including cerebrovascular reservoirs (e.g., Ommaya or Rikker reservoirs; see e.g., Raney, J. P. et al. (1988) *J. Neurosci. Nurs.* 20:23–29; Sundaresan, N. et al. (1989) *Oncology* 3:15–22), catheters for intrathecal delivery (e.g., Port-a-Cath, Y-catheters and the like; see e.g., Plummer, J. L. (1991) *Pain* 44:215–220; Yaksh, T. L. et al. (1986) *Pharmacol. Biochem. Behav.* 25:483–485), injectable intrathecal reservoirs (e.g., Spinalgesic; see e.g., Brazenor, G. A. (1987) *Neurosurgery* 21:484–491), implantable infusion pump systems (e.g., Infusaid; see e.g., Zierski, J. et al. (1988) *Acta Neurochem. Suppl.* 43:94–99; Kanoff, R. B. (1994) *J. Am. Osteopath. Assoc.* 94:487–493) and osmotic pumps (sold by Alza Corporation). A particularly preferred mode of administration is via an implantable, externally programmable infusion pump. Suitable infusion pump systems and reservoir systems are also described in U.S. Pat. No. 5,368,562 by Blomquist and U.S. Pat. No. 4,731,058 by Doan, developed by Pharmacia Deltec Inc.

The method of the invention for altering β-AP aggregation in vivo, and in particular for inhibiting β-AP aggregation, can be used therapeutically in diseases associated with abnormal β amyloid aggregation and deposition to thereby slow the rate of β amyloid deposition and/or lessen the degree of β amyloid deposition, thereby ameliorating the course of the disease. In a preferred embodiment, the method is used to treat Alzheimer's disease (e.g, sporadic or familial AD, including both individuals exhibiting symptoms of AD and individuals susceptible to familial AD). The method can also be used prophylactically or therapeutically to treat other clinical occurrences of, amyloid deposition, such as in Down's syndrome individuals and in patients with hereditary cerebral hemorrhage with amyloidosis-Dutch-type (HCHWA-D). While inhibition of β-AP aggregation is a preferred therapeutic method, modulators that promote β-AP aggregation may also be useful therapeutically by allowing for the sequestration of β-AP at sites that do not lead to neurological impairment.

Additionally, abnormal accumulation of β-amyloid precursor protein in muscle fibers has been implicated in the

pathology of sporadic inclusion body myositis (IBM) (Askanas, V. et al. (1996) *Proc. Natl. Acad Sci. USA* 93:1314–1319; Askanas, V. et al. (1995) *Current Opinion in Rheumatology* 7:486–496). Accordingly, the modulators of the invention can be used prophylactically or therapeutically in the treatment of disorders in which β-AP, or APP, is abnormally deposited at non-neurological locations, such as treatment of IBM by delivery of the modulators to muscle fibers.

VII. Unmodified Aβ Peptides that Inhibit Aggregation of Natural β-AP

In addition to the β-amyloid modulators described hereinbefore in which an Aβ peptide is coupled to a modifying group, the invention also provides β-amyloid modulators comprised of an unmodified Aβ peptide. It has now been discovered that certain portions of natural β-AP can alter aggregation of natural β-APs when contacted with the natural β-APs (see Example 12). Accordingly, these unmodified Aβ peptides comprise a portion of the natural β-AP sequence (i.e., a portion of βAP_{1-39}, βAP_{1-40}, βAP_{1-42} and βAP_{1-43}). In particular these unmodified Aβ peptides have at least one amino acid deletion compared to βAP_{1-39}, the shortest natural β-AP, such that the compound alters aggregation of natural β-amyloid peptides when contacted with the natural β-amyloid peptides. In various embodiments, these unmodified peptide compounds can promote aggregation of natural β-amyloid peptides, or, more preferably, can inhibit aggregation of natural β-amyloid peptides when contacted with the natural β-amyloid peptides. Even more preferably, the unmodified peptide compound inhibits aggregation of natural β-amyloid peptides when contacted with a molar excess amount of natural β-amyloid peptides (e.g., a 10-fold, 33-fold or 100-fold molar excess amount of natural β-AP).

As discussed above, the unmodified peptide compounds of the invention comprise an amino acid sequence having at least one amino acid deletion compared to the amino acid sequence of βAP_{1-39}. Alternatively, the unmodified peptide compound can have at least five, ten, fifteen, twenty, twenty-five, thirty or thirty-five amino acids deleted compared to βAP_{1-39}. Still further the unmodified peptide compound can have 1–5, 1–10, 1–15, 1–20, 1–25, 1–30 or 1–35 amino acids deleted compared to βAP_{1-39}. The amino acid deletion (s) may occur at the amino-terminus, the carboxy-terminus, an internal site, or a combination thereof, of the β-AP sequence. Accordingly, in one embodiment, an unmodified peptide compound of the invention comprises an amino acid sequence which has at least one internal amino acid deleted compared to βAP_{1-39}. Alternatively, the unmodified peptide compound can have at least five, ten, fifteen, twenty, twenty-five, thirty or thirty-five internal amino acids deleted compared to βAP_{1-39}. Still further the unmodified peptide compound can have 1–5, 1–10, 1–15, 1–20, 1–25, 1–30 or 1–35 internal amino acids deleted compared to βAP_{1-39}. For peptides with internal deletions, preferably the peptide has an amino terminus corresponding to amino acid residue 1 of natural βAP and a carboxy terminus corresponding to residue 40 of natural βAP and has one or more internal β-AP amino acid residues deleted (i.e., a non-contiguous Aβ peptide).

In another embodiment, the unmodified peptide compound comprises an amino acid sequence which has at least one N-terminal amino acid deleted compared to βAP_{1-39}. Alternatively, the unmodified peptide compound can have at least five, ten, fifteen, twenty, twenty-five, thirty or thirty-five N-terminal amino acids deleted compared to βAP_{1-39}. Still further the unmodified peptide compound can have 1–5, 1–10, 1–15, 1–20, 1–25, 1–30 or 1–35 N-terminal amino acids deleted compared to βAP_{1-39}.

In yet another embodiment, the unmodified peptide compound comprises an amino acid sequence which has at least one C-terminal amino acid deleted compared to βAP_{1-39} Alternatively, the unmodified peptide compound can have at least five, ten, fifteen, twenty, twenty-five, thirty or thirty-five C-terminal amino acids deleted compared to βAP_{1-39}. Still further the unmodified peptide compound can have 1–5, 1–10, 1–15, 1–20, 1–25, 1–30 or 1–35 C-terminal amino acids deleted compared to βAP_{1-39}.

In addition to deletion of amino acids as compared to βAP_{1-39}, the peptide compound can have additional non-β-AP amino acid residues added to it, for example, at the amino terminus, the carboxy-terminus or at an internal site. In one embodiment, the peptide compound has at least one non-β-amyloid peptide-derived amino acid at its N-terminus. Alternatively, the compound can have, for example, 1–3, 1–5, 1–7, 1–10, 1–15 or 1–20 non-β-amyloid peptide-derived amino acid at its N-terminus. In another embodiment, the peptide compound has at least one non-β-amyloid peptide-derived amino acid at its C-terminus. Alternatively, the compound can have, for example, 1–3, 1–5, 1–7, 1–10, 1–15 or 1–20 non-β-amyloid peptide-derived amino acid at its C-terminus.

In specific preferred embodiments, an unmodified peptide compound of the invention comprises $A\beta_{6-20}$ (the amino acid sequence of which is shown in SEQ ID NO:4), $A\beta_{16-30}$ (the amino acid sequence of which is shown in SEQ ID NO:14), $A\beta_{1-20, 26-40}$ (the amino acid sequence of which is shown in SEQ ID NO:15) or EEVVHHHHQQ-βAP_{16-40} (the amino acid sequence of which is shown in SEQ ID NO:16). In the nomenclature used herein, $\beta AP_{1-20, 26-40}$ represents βAP_{1-40} in which the internal amino acid residues 21–25 have been deleted.

An unmodified peptide compound of the invention can be chemically synthesized using standard techniques such as those described in Bodansky, M. *Principles of Peptide Synthesis*, Springer Verlag, Berlin (1993) and Grant, G. A. (ed.). *Synthetic Peptides: A User's Guide*, W. H. Freeman and Company, New York (1992). Automated peptide synthesizers are commercially available (e.g., Advanced ChemTech Model 396; Milligen/Biosearch 9600). Alternatively, unmodified peptide compounds can be prepared according to standard recombinant DNA techniques using a nucleic acid molecule encoding the peptide. A nucleotide sequence encoding the peptide can be determined using the genetic code and an oligonucleotide molecule having this nucleotide sequence can be synthesized by standard DNA synthesis methods (e.g., using an automated DNA synthesizer). Alternatively, a DNA molecule encoding an unmodified peptide compound can be derived from the natural β-amyloid precursor protein gene or cDNA (e.g., using the polymerase chain reaction and/or restriction enzyme digestion) according to standard molecular biology techniques.

Accordingly, the invention further provides an isolated nucleic acid molecule comprising a nucleotide sequence encoding a β-amyloid peptide compound, the β-amyloid peptide compound comprising an amino acid sequence having at least one amino acid deletion compared to βAP_{1-39} such that the β-amyloid peptide compound alters aggregation of natural β-amyloid peptides when contacted with the natural β-amyloid peptides. As used herein, the term "nucleic acid molecule" is intended to include DNA molecules and RNA molecules and may be single-stranded or double-stranded, but preferably is double-stranded DNA.

The isolated nucleic acid encodes a peptide wherein one or more amino acids are deleted from the N-terminus, C-terminus and/or an internal site of βAP_{1-39}, as discussed above. In yet other embodiments, the isolated nucleic acid encodes a peptide compound having one or more amino acids deleted compared to βAP_{1-39} and further having at least one non-β-AP derived amino acid residue added to it, for example, at the amino terminus, the carboxy-terminus or at an internal site. In specific preferred embodiments, an isolated nucleic acid molecule of the invention encodes βAP_{6-20}, βAP_{16-30}, $\beta AP_{1-20, 26-40}$ or EEVVHHHHQQ-βAP_{16-40} (SEQ ID NO:16).

To facilitate expression of a peptide compound in a host cell by standard recombinant DNA techniques, the isolated nucleic acid encoding the peptide is incorporated into a recombinant expression vector. Accordingly, the invention also provides recombinant expression vectors comprising the nucleic acid molecules of the invention. As used herein, the term "vector" refers to a nucleic acid molecule capable of transporting another nucleic acid to which it has been linked. One type of vector is a "plasmid", which refers to a circular double stranded DNA loop into which additional DNA segments may be ligated. Another type of vector is a viral vector, wherein additional DNA segments may be ligated into the viral genome. Certain vectors are capable of autonomous replication in a host cell into which they are introduced (e.g., bacterial vectors having a bacterial origin of replication and episomal mammalian vectors). Other vectors (e.g., non-episomal mammalian vectors) are integrated into the genome of a host cell upon introduction into the host cell, and thereby are replicated along with the host genome. Moreover, certain vectors are capable of directing the expression of genes to which they are operatively linked. Such vectors are referred to herein as "recombinant expression vectors" or simply "expression vectors". In general, expression vectors of utility in recombinant DNA techniques are often in the form of plasmids. In the present specification, "plasmid" and "vector" may be used interchangeably as the plasmid is the most commonly used form of vector. However, the invention is intended to include such other forms of expression vectors, such as viral vectors, which serve equivalent functions.

In the recombinant expression vectors of the invention, the nucleotide sequence encoding the peptide compound are operatively linked to one or more regulatory sequences, selected on the basis of the host cells to be used for expression. The term "operably linked" is intended to mean that the sequences encoding the peptide compound are linked to the regulatory sequence(s) in a manner that allows for expression of the peptide compound. The term "regulatory sequence" is intended to includes promoters, enhancers and other expression control elements (e.g., polyadenylation signals). Such regulatory sequences are described, for example, in Goeddel; *Gene Expression Technology: Methods in Enzymology* 185, Academic Press, San Diego, Calif. (1990). Regulatory sequences include those that direct constitutive expression of a nucleotide sequence in many types of host cell, those that direct expression of the nucleotide sequence only in certain host cells (e.g., tissue-specific regulatory sequences) and those that direct expression in a regulatable manner (e.g., only in the presence of an inducing agent). It will be appreciated by those skilled in the art that the design of the expression vector may depend on such factors as the choice of the host cell to be transformed, the level of expression of peptide compound desired, etc. The expression vectors of the invention can be introduced into host cells thereby to produce peptide compounds encoded by nucleic acids as described herein.

The recombinant expression vectors of the invention can be designed for expression of peptide compounds in prokaryotic or eukaryotic cells. For example, peptide compounds can be expressed in bacterial cells such as *E. coli*, insect cells (using baculovirus expression vectors) yeast cells or mammalian cells. Suitable host cells are discussed further in Goeddel, *Gene Expression Technology: Methods in Enzymology* 185, Academic Press, San Diego, Calif. (1990). Alternatively, the recombinant expression vector may be transcribed and translated in vitro, for example using T7 promoter regulatory sequences and T7 polymerase. Examples of vectors for expression in yeast *S. cerivisae* include pYepSec1 (Baldari et al., (1987) *EMBO J.* 6:229–234), pMFa (Kurjan and Herskowitz, (1982) *Cell* 30:933–943), pJRY88 (Schultz et al., (1987) *Gene* 54:113–123), and pYES2 (Invitrogen Corporation, San Diego, Calif.). Baculovirus vectors available for expression of proteins or peptides in cultured insect cells (e.g., Sf 9 cells) include the pAc series (Smith et al., (1983) *Mol. Cell. Biol.* 3:2156–2165) and the pVL series (Lucklow, V. A., and Summers, M. D., (1989) *Virology* 170:31–39). Examples of mammalian expression vectors include pCDM8 (Seed, B., (1987) *Nature* 329:840) and pMT2PC (Kaufman et al. (1987), *EMBO J.* 6:187–195). When used in mammalian cells, the expression vector's control functions are often provided by viral regulatory elements. For example, commonly used promoters are derived from polyoma, Adenovirus 2, cytomegalovirus and Simian Virus 40.

In addition to the regulatory control sequences discussed above, the recombinant expression vector may contain additional nucleotide sequences. For example, the recombinant expression vector may encode a selectable marker gene to identify host cells that have incorporated the vector. Such selectable marker genes are well known in the art. Moreover, the facilitate secretion of the peptide compound from a host cell, in particular mammalian host cells, the recombinant expression vector preferably encodes a signal sequence operatively linked to sequences encoding the amino-terminus of the peptide compound such that upon expression, the peptide compound is synthesized with the signal sequence fused to its amino terminus. This signal sequence directs the peptide compound into the secretory pathway of the cell and is then cleaved, allowing for release of the mature peptide compound (i.e., the peptide compound without the signal sequence) from the host cell. Use of a signal sequence to facilitate secretion of proteins or peptides from mammalian host cells is well known in the art.

A recombinant expression vector comprising a nucleic acid encoding a peptide compound that alters aggregation of natural β-AP can be introduced into a host cell to thereby produce the peptide compound in the host cell. Accordingly, the invention also provides host cells containing the recombinant expression vectors of the invention. The terms "host cell" and "recombinant host cell" are used interchangeably herein. It is understood that such terms refer not only to the particular subject cell but to the progeny or potential progeny of such a cell. Because certain modifications may occur in succeeding generations due to either mutation or environmental influences, such progeny may not, in fact, be identical to the parent cell, but are still included within the scope of the term as used herein. A host cell may be any prokaryotic or eukaryotic cell. For example, a peptide compound may be expressed in bacterial cells such as *E. coli*, insect cells, yeast or mammalian cells. Preferably, the peptide compound is expressed in mammalian cells. In a preferred embodiment, the peptide compound is expressed in mammalian cells in vivo in a mammalian subject to treat

amyloidosis in the subject through gene therapy (discussed further below). Preferably, the β-amyloid peptide compound encoded by the recombinant expression vector is secreted from the host cell upon being expressed in the host cell.

Vector DNA can be introduced into prokaryotic or eukaryotic cells via conventional transformation or transfection techniques. As used herein, the terms "transformation" and "transfection" are intended to refer to a variety of art-recognized techniques for introducing foreign nucleic acid (e.g., DNA) into a host cell, including calcium phosphate or calcium chloride co-precipitation, DEAE-dextran-mediated transfection, lipofection, electroporation, microinjection and viral-mediated transfection. Suitable methods for transforming or transfecting host cells can be found in Sambrook et al. (*Molecular Cloning: A Laboratory Manual,* 2nd Edition, Cold Spring Harbor Laboratory press (1989)), and other laboratory manuals. Methods for introducing DNA into mammalian cells in vivo are also known in the art and can be used to deliver the vector DNA to a subject for gene therapy purposes (discussed further below).

For stable transfection of mammalian cells, it is known that, depending upon the expression vector and transfection technique used, only a small fraction of cells may integrate the foreign DNA into their genome. In order to identify and select these integrants, a gene that encodes a selectable marker (e.g., resistance to antibiotics) is generally introduced into the host cells along with the gene of interest. Preferred selectable markers include those that confer resistance to drugs, such as G418, hygromycin and methotrexate. Nucleic acid encoding a selectable marker may be introduced into a host cell on the same vector as that encoding the peptide compound or may be introduced on a separate vector. Cells stably transfected with the introduced nucleic acid can be identified by drug selection (e.g., cells that have incorporated the selectable marker gene will survive, while the other cells die).

A nucleic acid of the invention can be delivered to cells in vivo using methods known in the art, such as direct injection of DNA, receptor-mediated DNA uptake or viral-mediated transfection. Direct injection has been used to introduce naked DNA into cells in vivo (see e.g., Acsadi et al. (1991) *Nature* 332:815–818; Wolff et al. (1990) *Science* 247:1465–1468). A delivery apparatus (e.g., a "gene gun") for injecting DNA into cells in vivo can be used. Such an apparatus is commercially available (e.g., from BioRad). Naked DNA can also be introduced into cells by complexing the DNA to a cation, such as polylysine, which is coupled to a ligand for a cell-surface receptor (see for example Wu, G. and Wu, C. H. (1988) *J. Biol. Chem.* 263:14621; Wilson el al. (1992) *J. Biol. Chem.* 267:963–967; and U.S. Pat. No. 5,166,320). Binding of the DNA-ligand complex to the receptor facilitates uptake of the DNA by receptor-mediated endocytosis. Additionally, a DNA-ligand complex linked to adenovirus capsids which naturally disrupt endosomes, thereby releasing material into the cytoplasm can be used to avoid degradation of the complex by intracellular lysosomes (see for example Curiel el al. (1991) *Proc. Natl. Acad. Sci. USA* 88:8850; Cristiano et al. (1993) *Proc. Natl. Acad. Sci. USA* 90:2122–2126).

Defective retroviruses are well characterized for use in gene transfer for gene therapy purposes (for a review see Miller, A. D. (1990) *Blood* 76:271). Protocols for producing recombinant retroviruses and for infecting cells in vitro or in vivo with such viruses can be found in *Current Protocols in Molecular Biology*, Ausubel, F. M. et al. (eds.) Greene Publishing Associates, (1989), Sections 9.10–9.14 and other standard laboratory manuals. Examples of suitable retroviruses include pLJ, pZIP, pWE and pEM which are well known to those skilled in the art. Examples of suitable packaging virus lines include ψCrip, ψCre, ψ2 and ψAm. Retroviruses have been used to introduce a variety of genes into many different cell types, including epithelial cells, endothelial cells, lymphocytes, myoblasts, hepatocytes, bone marrow cells, in vitro and/or in vivo (see for example Eglitis, et al. (1985) *Science* 230:1395–1398; Danos and Mulligan (1988) *Proc. Natl. Acad. Sci. USA* 85:6460–6464; Wilson et al. (1988) *Proc. Natl. Acad. Sci. USA* 85:3014–3018; Armentano et al. (1990) *Proc. Natl. Acad. Sci. USA* 87:6141–6145; Huber et al. (1991) *Proc. Natl. Acad. Sci. USA* 88:8039–8043; Ferry et al. (1991) *Proc. Natl. Acad. Sci. USA* 88:8377–8381; Chowdhury et al. (1991) *Science* 254:1802–1805; van Beusechem et al. (1992) *Proc. Natl. Acad. Sci. USA* 89:7640–7644; Kay et al. (1992) *Human Gene Therapy* 3:641–647; Dai et al. (1992) *Proc. Natl. Acad. Sci. USA* 89:10892–10895; Hwu et al. (1993) *J. Immunol.* 150:4104–4115; U.S. Pat. No. 4,868,116; U.S. Pat. No. 4,980,286; PCT Application WO 89/07136; PCT Application WO 89/02468; PCT Application WO 89/05345; and PCT Application WO 92/07573).

Alternatively, the genome of an adenovirus can be manipulated such that it encodes and expresses a peptide compound but is inactivated in terms of its ability to replicate in a normal lytic viral life cycle. See for example Berkner et al. (1988) *BioTechniques* 6:616; Rosenfeld et al. (1991) *Science* 252:431–434; and Rosenfeld et al. (1992) *Cell* 68:143–155. Suitable adenoviral vectors derived from the adenovirus strain Ad type 5 dl324 or other strains of adenovirus (e.g., Ad2, Ad3, Ad7 etc.) are well known to those skilled in the art. Recombinant adenoviruses are advantageous in that they do not require dividing cells to be effective gene delivery vehicles and can be used to infect a wide variety of cell types, including airway epithelium (Rosenfeld et al. (1992) cited supra), endothelial cells (Lemarchand et al. (1992) *Proc. Natl. Acad. Sci. USA* 89:6482–6486), hepatocytes (Herz and Gerard (1993) *Proc. Natl. Acad. Sci. USA* 90:2812–2816) and muscle cells (Quantin el al. (1992) *Proc. Natl. Acad. Sci. USA* 89:2581–2584). Additionally, introduced adenoviral DNA (and foreign DNA contained therein) is not integrated into the genome of a host cell but remains episomal, thereby avoiding potential problems that can occur as a result of insertional mutagenesis in situations where introduced DNA becomes integrated into the host genome (e.g., retroviral DNA).

Adeno-associated virus (AAV) can also be used for delivery of DNA for gene therapy purposes. AAV is a naturally occurring defective virus that requires another virus, such as an adenovirus or a herpes virus, as a helper virus for efficient replication and a productive life cycle. (For a review see Muzyczka et al. *Curr. Topics in Micro. and Immunol.* (1992) 158:97–129). It is also one of the few viruses that may integrate its DNA into non-dividing cells, and exhibits a high frequency of stable integration (see for example Flotte et al. (1992) *Am. J. Respir. Cell. Mol. Biol.* 7:349–356; Samulski et al. (1989) *J. Virol.* 63:3822–3828; and McLaughlin et al. (1989) *J. Virol.* 62:1963–1973). Vectors containing as little as 300 base pairs of AAV can be packaged and can integrate. An AAV vector such as that described in Tratschin et al. (1985) *Mol. Cell. Biol.* 5:3251–3260 can be used to introduce DNA into cells. A variety of nucleic acids have been introduced into different cell types using AAV vectors (see for example Hermonat et al. (1984) *Proc. Natl. Acad. Sci. USA* 81:6466–6470; Tratschin et al. (1985) *Mol. Cell. Biol.* 4:2072–2081; Wondisford et al. (1988) *Mol. Endocrinol.*

2:32–39; Tratschin et al. (1984) *J. Virol.* 51:611–619; and Flotte et al. (1993) *J. Biol. Chem.* 268:3781–3790).

The invention provides a method for treating a subject for a disorder associated with β-amyloidosis, comprising administering to the subject a recombinant expression vector encoding a β-amyloid peptide compound, the compound comprising an amino acid sequence having at least one amino acid deletion compared to βAP_{1-39}, such that the β-amyloid peptide compound is synthesized in the subject and the subject is treated for a disorder associated with β-amyloidosis. Preferably, the disorder is Alzheimer's disease. In one embodiment the recombinant expression vector directs expression of the peptide compound in neuronal cells. In another embodiment, the recombinant expression vector directs expression of the peptide compound in glial cells. In yet another embodiment, the recombinant expression vector directs expression of the peptide compound in fibroblast cells.

General methods for gene therapy are known in the art. See for example, U.S. Pat. No. 5,399,346 by Anderson et al. A biocompatible capsule for delivering genetic material is described in PCT Publication WO 95/05452 by Baetge et al. Methods for grafting genetically modified cells to treat central nervous system disorders are described in U.S. Pat. No. 5,082,670 and in PCT Publications WO 90/06757 and WO 93/10234, all by Gage et al. Isolation and/or genetic modification of multipotent neural stem cells or neuro-derived fetal cells are described in PCT Publications WO 94/02593 by Anderson et al., WO 94/16718 by Weiss et al., and WO 94/23754 by Major et al. Fibroblasts transduced with genetic material are described in PCT Publication WO 89/02468 by Mulligan et al. Adenovirus vectors for transfering genetic material into cells of the central nervous system are described in PCT Publication WO 94/08026 by Kahn et al. Herpes simplex virus vectors suitable for treating neural disorders are described in PCT Publications WO 94/04695 by Kaplitt and WO 90/09441 by Geller et al. Promoter elements of the glial fibrillary acidic protein that can confer astrocyte specific expression on a linked gene or gene fragment, and which thus can be used for expression of Aβ peptides specifically in astrocytes, is described in PCT Publication WO 93/07280 by Brenner et al. Furthermore, alternative to expression of an Aβ peptide to modulate amyloidosis, an antisense oligonucleotide that is complementary to a region of the β-amyloid precursor protein mRNA corresponding to the peptides described herein can be expressed in a subject to modulate amyloidosis. General methods for expressing antisense oligonucleotides to modulate nervous system disorders are described in PCT Publication WO 95/09236.

Alternative to delivery by gene therapy, a peptide compound of the invention comprising an amino acid sequence having at least one amino acid deletion compared to βAP_{1-39} can be delivered to a subject by directly administering the peptide compound to the subject as described further herein for the modified peptide compounds of the invention. The peptide compound can be formulated into a pharmaceutical composition comprising a therapeutically effective amount of the β-amyloid peptide compound and a pharmaceutically acceptable carrier. The peptide compound can be contacted with natural β-amyloid peptides with a β-amyloid peptide compound such that aggregation of the natural β-amyloid peptides is inhibited. Moreover, the peptide compound can be administered to the subject in a therapeutically effective amount such that the subject is treated for a disorder associated with β-amyloidosis, such as Alzheimer's disease.

VIII. Other Embodiments

Although the invention has been illustrated hereinbefore with regard to Aβ peptide compounds, the principles described, involving attachment of a modifying group(s) to a peptide compound, are applicable to any amyloidogenic protein or peptide as a means to create a modulator compound that modulates, and preferably inhibits, amyloid aggregation. Accordingly, the invention provides modulator compounds that can be used to treat amyloidosis in a variety of forms and clinical settings.

Amyloidosis is a general term used to describe pathological conditions characterized by the presence of amyloid. Amyloid is a general term referring to a group of diverse but specific extracellular protein deposits which are seen in a number of different diseases. Though diverse in their occurrence, all amyloid deposits have common morphologic properties, stain with specific dyes (e.g., Congo red), and have a characteristic red-green birefringent appearance in polarized light after staining. They also share common ultrastructural features and common x-ray diffraction and infrared spectra. Amyloidosis can be classified clinically as primary, secondary, familial and/or isolated. Primary amyloid appears de novo without any preceding disorder. Secondary amyloid is that form which appears as a complication of a previously existing disorder. Familial amyloid is a genetically inherited form found in particular geographic populations. Isolated forms of amyloid are those that tend to involve a single organ system.

Different amyloids are characterized by the type of protein(s) or peptide(s) present in the deposit. For example, as described hereinbefore, amyloid deposits associated with Alzheimer's disease comprise the β-amyloid peptide and thus a modulator compound of the invention for detecting and/or treating Alzheimer's disease is designed based on modification of the β-amyloid peptide. The identities of the protein(s) or peptide(s) present in amyloid deposits associated with a number of other amyloidogenic diseases have been elucidated. Accordingly, modulator compounds for use in the detection and/or treatment of these other amyloidogenic diseases can be prepared in a similar fashion to that described herein for β-AP-derived modulators. In vitro assay systems can be established using an amyloidogenic protein or peptide which forms fibrils in vitro, analogous to the Aβ assays described herein. Modulators can be identified using such assay systems, based on the ability of the modulator to disrupt the β-sheet structure of the fibrils. Initially, an entire amyloidogenic protein can be modified or, more preferably, a peptide fragment thereof that is known to form fibrils in vitro can be modified (e.g., analogous to Aβ1–40 described herein). Amino acid deletion and substitution analyses can then be performed on the modified protein or peptide (analogous to the studies described in the Examples) to delineate an aggregation core domain that is sufficient, when modified, to disrupt fibril formation.

Non-limiting examples of amyloidogenic proteins or peptides, and their associated amyloidogenic disorders, include:

Transthyretin (TTR)—Amyloids containing transthyretin occur in familial amyloid polyneuropathy (Portuguese, Japanese and Swedish types), familial amyloid cardiomyopathy (Danish type), isolated cardiac amyloid and systemic senile amyloidosis. Peptide fragments of transthyretin have been shown to form amyloid fibrils in vitro. For example, TTR 10–20 and TTR 105–115 form amyloid-like fibrils in 20–30% acetonitrile/water at room temperature (Jarvis, J. A., et al. (1994) *Int. J. Pept. Protein Res.* 44:388–398). Moreover, familial cardiomyopathy (Danish type) is asso-

ciated with mutation of Leu at position 111 to Met, and an analogue of TTR 105–115 in which the wildtype Leu at position 111 has been substituted with Met (TTR 105–115Met111) also forms amyloid-like fibrils in vitro (see e.g., Hermansen, L. F., et al. (1995) *Eur. J. Biochem.* 227:772–779; Jarvis et al. supra). Peptide fragments of TTR that form amyloid fibrils in vitro are also described in Jarvis, J. A., et al. (1993) *Biochem. Biophys. Res. Commun.* 192:991–998 and Gustavsson, A., et al. (1991) *Biochem. Biophys. Res. Commun.* 175:1159–1164. A peptide fragment of wildtype or mutated transthyretin that forms amyloid fibrils can be modified as described herein to create a modulator of amyloidosis that can be used in the detection or treatment of familial amyloid polyneuropathy (Portuguese, Japanese and Swedish types), familial amyloid cardiomyopathy (Danish type), isolated cardiac amyloid or systemic senile amyloidosis.

Prion Protein (PrP)—Amyloids in a number of spongiform encephalopathies, including scrapie in sheep, bovine spongiform encephalopathy in cows and Creutzfeldt-Jakob disease (CJ) and Gerstmann-Straussler-Scheinker syndrome (GSS) in humans, contain PrP. Limited proteolysis of PrPSc (the prion protein associated with scrapie) leads to a 27–30 kDa fragment (PrP27–30) that polymerizes into rod-shaped amyloids (see e.g., Pan, K. M., et al. (1993) *Proc. Natl. Acad. Sci. USA* 90:10962–10966; Gasset, M., et al. (1993) *Proc. Natl. Acad. Sci. USA* 90:1–5). Peptide fragments of PrP from humans and other mammals have been shown to form amyloid fibrils in vitro. For example, polypeptides corresponding to sequences encoded by normal and mutant alleles of the PRNP gene (encoding the precursor of the prion protein involved in CJ), in the regions of codon 178 and codon 200, spontaneously form amyloid fibrils in vitro (see e.g., Goldfarb, L. G., et al. (1993) *Proc. Natl. Acad. Sci. USA* 90:4451–4454). A peptide encompassing residues 106–126 of human PrP has been reported to form straight fibrils similar to those extracted from GSS brains, whereas a peptide encompassing residues 127–147 of human PrP has been reported to form twisted fibrils resembling scrapie-associated fibrils (Tagliavini, F., et al. (1993) *Proc. Natl. Acad. Sci. USA* 90:9678–9682). Peptides of Syrian hamster PrP encompassing residues 109–122, 113–127, 113–120, 178–191 or 202–218 have been reported to form amyloid fibrils, with the most amyloidogenic peptide being Ala-Gly-Ala-Ala-Ala-Ala-Gly-Ala (SEQ ID NO:17), which corresponds to residues 113–120 of Syrian hamster PrP but which is also conserved in PrP from other species (Gasset, M., et al. (1992) *Proc. Natl. Acad. Sci. USA* 89:10940–10944). A peptide fragment of PrP that forms amyloid fibrils can be modified as described herein to create a modulator of amyloidosis that can be used in the detection or treatment of scrapie, bovine spongiform encephalopathy, Creutzfeldt-Jakob disease or Gerstmann-Straussler-Scheinker syndrome.

Islet Amyloid Polypeptide (IAPP, also known as amylin)—Amyloids containing IAPP occur in adult onset diabetes and insulinoma. IAPP is a 37 amino acid polypeptide formed from an 89 amino acid precursor protein (see e.g., Betsholtz, C., et al. (1989) *Exp. Cell. Res.* 183:484–493; Westermark, P., et al. (1987) *Proc. Natl. Acad. Sci. USA* 84:3881–3885). A peptide corresponding to IAPP residues 20–29 has been reported to form amyloid-like fibrils in vitro, with residues 25–29, having the sequence Ala-Ile-Leu-Ser-Ser (SEQ ID NO:18), being strongly amyloidogenic (Westermark, P., et al. (1990) *Proc. Natl. Acad. Sci. USA* 87:5036–5040; Glenner, G. G., et al. (1988) *Biochem. Biophys. Res. Commun.* 155:608–614). A peptide fragment of IAPP that forms amyloid fibrils can be modified as described herein to create a modulator of amyloidosis that can be used in the detection or treatment of adult onset diabetes or insulinoma.

Atrial Natriuretic Factor (ANF)—Amyloids containing ANF are associated with isolated atrial amyloid (see e.g., Johansson, B., et al. (1987) *Biochem. Biophys. Res. Commun.* 148:1087–1092). ANF corresponds to amino acid residues 99–126 (proANF99–126) of the ANF prohormone (proANP1–126) (Pucci, A., et al. (1991) *J. Pathol.* 165:235–241). ANF, or a fragment thereof, that forms amyloid fibrils can be modified as described herein to create a modulator of amyloidosis that can be used in the detection or treatment of isolated atrial amyloid.

Kappa or Lambda Light Chain—Amyloids containing kappa or lambda light chains are associated idiopathic (primary) amyloidosis, myeloma or macroglobulinemia-associated amyloidosis, and primary localized cutaneous nodular amyloidosis associated with Sjogren's syndrome. The structure of amyloidogenic kappa and lambda light chains, including amino acid sequence analysis, has been characterized (see e.g., Buxbaum, J. N., et al. (1990) *Ann. Intern. Med.* 112:455–464; Schormann, N., et al. (1995) *Proc. Natl. Acad. Sci. USA* 92:9490–9494; Hurle, M. R., et al. (1994) *Proc. Natl. Acad. Sci. USA* 91:5446–5450; Liepnieks, J. J., et al. (1990) *Mol. Immunol.* 27:481–485; Gertz, M. A., et al. (1985) *Scand. J. Immunol.* 22:245–250; Inazumi, T., et al. (1994) *Dermatology* 189:125–128). Kappa or lambda light chains, or a peptide fragment thereof that forms amyloid fibrils, can be modified as described herein to create a modulator of amyloidosis that can be used in the detection or treatment of idiopathic (primary) amyloidosis, myeloma or macroglobulinemia-associated amyloidosis or primary localized cutaneous nodular amyloidosis associated with Sjogren's syndrome.

Amyloid A—Amyloids containing the amyloid A protein (AA protein), derived from serum amyloid A, are associated with reactive (secondary) amyloidosis (see e.g., Liepnieks, J. J., et al. (1995) *Biochim. Biophys. Acta* 1270:81–86), familial Mediterranean Fever and familial amyloid nephropathy with urticaria and deafness (Muckle-Wells syndrome) (see e.g., Linke, R. P., et al. (1983) *Lab. Invest.* 48:698–704). Recombinant human serum amyloid A forms amyloid-like fibrils in vitro (Yamada, T., et al. (1994) *Biochim. Biophys. Acta* 1226:323–329) and circular dichroism studies revealed a predominant β sheet/turn structure (McCubbin, W. D., et al. (1988) *Biochem J.* 256:775–783). Serum amyloid A, amyloid A protein or a fragment thereof that forms amyloid fibrils can be modified as described herein to create a modulator of amyloidosis that can be used in the detection or treatment of reactive (secondary) amyloidosis, familial Mediterranean Fever and familial amyloid nephropathy with urticaria and deafness (Muckle-Wells syndrome).

Cystatin C—Amyloids containing a variant of cystatin C are associated with hereditary cerebral hemorrhage with amyloidosis of Icelandic type. The disease is associated with a leucine to glycine mutation at position 68 and cystatin C containing this mutation aggregates in vitro (Abrahamson, M. and Grubb, A. (1994) *Proc. Natl. Acad. Sci. USA* 91:1416–1420). Cystatin C or a peptide fragment thereof that forms amyloid fibrils can be modified as described herein to create a modulator of amyloidosis that can be used in the detection or treatment of hereditary cerebral hemorrhage with amyloidosis of Icelandic type.

β2 microglobulin—Amyloids containing β2 microglobulin (β2M) are a major complication of long term hemodialysis (see e.g., Stein, G., et al. (1994) Nephrol. Dial.

Transplant. 9:48–50; Floege, J., et al. (1992) Kidney Int. Suppl. 38:S78–S85; Maury, C. P. (1990) Rheumatol. Int. 10:1–8). The native β2M protein has been shown to form amyloid fibrils in vitro (Connors, L. H., et al. (1985) Biochem. Biophys. Res. Commun. 131:1063–1068; Ono, K., et al. (1994) Nephron 66:404–407). β2M, or a peptide fragment thereof that forms amyloid fibrils, can be modified as described herein to create a modulator of amyloidosis that can be used in the detection or treatment of amyloidosis associated with long term hemodialysis.

Apolipoprotein A-I (ApoA-I)—Amyloids containing variant forms of ApoA-I have been found in hereditary non-neuropathic systemic amyloidosis (familial amyloid polyneuropathy III). For example, N-terminal fragments (residues 1–86, 1–92 and 1–93) of an ApoA-I variant having a Trp to Arg mutation at position 50 have been detected in amyloids (Booth, D. R., et al. (1995) *QJM* 88:695–702). In another family, a leucine to arginine mutation at position 60 was found (Soutar, A. K., et al. (1992) *Proc. Natl. Acad. Sci. USA* 89:7389–7393). ApoA-I or a peptide fragment thereof that forms amyloid fibrils can be modified as described herein to create a modulator of amyloidosis that can be used in the detection or treatment of hereditary non-neuropathic systemic amyloidosis.

Gelsolin—Amyloids containing variants of gelsolin are associated with familial amyloidosis of Finnish type. Synthetic gelsolin peptides that have sequence homology to wildtype or mutant gelsolins and that form amyloid fibrils in vitro are reported in Maury, C. P. et al. (1994) *Lab. Invest.* 70:558–564. A nine residue segment surrounding residue 187 (which is mutated in familial gelsolin amyloidosis) was defined as an amyloidogenic region (Maury, et al., supra; see also Maury, C. P., et al. (1992) *Biochem. Biophys. Res. Commun.* 183:227–231; Maury, C. P. (1991) *J. Clin. Invest.* 87:1195–1199). Gelsolin or a peptide fragment thereof that forms amyloid fibrils can be modified as described herein to create a modulator of amyloidosis that can be used in the detection or treatment of familial amyloidosis of Finnish type.

Procalcitonin or calcitonin—Amyloids containing procalcitonin, calcitonin or calcitonin-like immunoreactivity have been detected in amyloid fibrils associated with medullary carcinoma of the thyroid (see e.g., Butler, M. and Khan, S. (1986) Arch. Pathol. Lab. Med. 110:647–649; Sletten, K., et al. (1976) J. Exp. Med. 143:993–998). Calcitonin has been shown to form a nonbranching fibrillar structure in vitro (Kedar, I., et al. (1976) Isr. J. Med. Sci. 12:1137–1140). Procalcitonin, calcitonin or a fragment thereof that forms amyloid fibrils can be modified as described herein to create a modulator of amyloidosis that can be used in the detection or treatment of amyloidosis associated with medullary carcinoma of the thyroid.

Fibrinogen—Amyloids containing a variant form of fibrinogen alpha-chain have been found in hereditary renal amyloidosis. An arginine to leucine mutation at position 554 has been reported in amyloid fibril protein isolated from postmortem kidney of an affected individual (Benson, M. D., et al. (1993) *Nature Genetics* 3:252–255). Fibrinogen alpha-chain or a peptide fragment thereof that forms amyloid fibrils can be modified as described herein to create a modulator of amyloidosis that can be used in the detection or treatment of fibrinogen-associated hereditary renal amyloidosis.

Lysozyme—Amyloids containing a variant form of lysozyme have been found in hereditary systemic amyloidosis. In one family the disease was associated with a threonine to isoleucine mutation at position 56, whereas in another family the disease was associated with a histidine to aspartic acid mutation at position 67 (Pepys, M. B., et al. (1993) *Nature* 362:553–557). Lysozyme or a peptide fragment thereof that forms amyloid fibrils can be modified as described herein to create a modulator of amyloidosis that can be used in the detection or treatment of lysozyme-associated hereditary systemic amyloidosis.

This invention is further illustrated by the following examples which should not be construed as limiting. A modulator's ability to alter the aggregation of β-amyloid peptide in the assays described below are predictive of the modulator's ability to perform the same function in vivo. The contents of all references, patents and published patent applications cited throughout this application are hereby incorporated by reference.

EXAMPLE 1

Construction of β-Amyloid Modulators

A β-amyloid modulator composed of an amino-terminally biotinylated β-amyloid peptide of the amino acid sequence: DAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVV (positions 1 to 40 of SEQ ID NO:1) was prepared by solid-phase peptide synthesis using an N^{α}-9-fluorenylmethyloxycarbonyl (FMOC)-based protection strategy as follows. Starting with 2.5 mmoles of FMOC-Val-Wang resin, sequential additions of each amino acid were performed using a four-fold excess of protected amino acids, 1-hydroxybenzotriazole (HOBt) and diisopropyl carbodiimide (DIC). Recouplings were performed when necessary as determined by ninhydrin testing of the resin after coupling. Each synthesis cycle was minimally described by a three minute deprotection (25% piperidine/N-methyl-pyrrolidone (NMP)), a 15 minute deprotection, five one minute NMP washes, a 60 minute coupling cycle, five NMP washes and a ninhydrin test. To a 700 mg portion of the fully assembled peptide-resin, biotin (obtained commercially from Molecular Probes, Inc.) was substituted for an FMOC-amino acid was coupled by the above protocol. The peptide was removed from the resin by treatment with trifluoroacetic acid (TFA) (82.5%), water (5%), thioanisole (5%), phenol (5%), ethanedithiol (2.5%) for two hours followed by precipitation of the peptide in cold ether. The solid was pelleted by centrifugation (2400 rpm×10 min.), and the ether decanted. It was resuspended in ether, pelleted and decanted a second time. The solid was dissolved in 10% acetic acid and lyophilized to dryness to yield 230 mg of crude biotinylated peptide. 60 mg of the solid was dissolved in 25% acetonitrile (ACN)/0.1% TFA and applied to a C18 reversed phase high performance liquid chromatography (HPLC) column. Biotinyl βAP_{1-40} was eluted using a linear gradient of 30–45% acetonitrile/0.1% TFA over 40 minutes. One primary fraction (4 mg) and several side fractions were isolated. The main fraction yielded a mass spectrum of 4556 (matrix-assisted laser desorption ionization-time of flight) which matches the theoretical (4555) for this peptide.

A β-amyloid modulator composed of an amino-terminally biotinylated β-amyloid peptide of the amino acid sequence: DAEFRHDSGYEVHHQ (positions 1 to 15 of SEQ ID NO:1) was prepared on an Advanced ChemTech Model 396 multiple peptide synthesizer using an automated protocol established by the manufacturer for 0.025 mmole scale synthesis. Double couplings were per-

formed on all cycles using 2-(1H-benzotriazol-1-yl)-1, 1,3,3-tetramethyluronium hexafluorophosphate (HBTU)/N,N-diisopropylethylamine (DIEA)/HOBt/ FMOC-AA in four-fold excess for 30 minutes followed by DIC/HOBt/FMOC-AA in four-fold excess for 45 minutes. The peptide was deprotected and removed from the resin by treatment with TFA/water (95%/5%) for three hours and precipitated with ether as described above. The pellet was resuspended in 10% acetic acid and lyophilized. The material was purified by a preparative HPLC using 15%–40% acetonitrile over 80 minutes on a Vydac C18 column (21×250 mm). The main isolate eluted as a single symmetrical peak when analyzed by analytical HPLC and yielded the expected molecular weight when analyzed by electrospray mass spectrometry. Result=2052.6 (2052 theoretical).

β-amyloid modulator compounds comprising other regions of the β-AP amino acid sequence (e.g., an Aβ aggregation core domain) were similarly prepared using the synthesis methods described above. Moreover, modulators comprising other amyloidogenic peptides can be similarly prepared.

EXAMPLE 2

Inhibition of β-Amyloid Aggregation by Modulators

The ability of β-amyloid modulators to inhibit the aggregation of natural β-AP when combined with the natural β-AP was examined in a series of aggregation assays. Natural β-AP (β-AP_{1-40}) was obtained commercially from Bachem (Torrance, Calif.). Amino-terminally biotinylated β-AP modulators were prepared as described in Example 1.

A. Optical Density Assay

In one assay, β-AP aggregation was measured by determining the increase in turbidity of a solution of natural β-AP over time in the absence or presence of various concentrations of the modulator. Turbidity of the solution was quantitated by determining the optical density at 400 nm ($A_{400\ nm}$) of the solution over time.

The aggregation of natural β-AP in the absence of modulator was determined as follows. β-AP_{1-40} was dissolved in hexafluoro isopropanol (HFIP; Aldrich Chemical Co., Inc.) at 2 mg/ml. Aliquots of the HFIP solution (87 μl) were transferred to individual 10 mm×75 mm test tubes. A stream of argon gas was passed through each tube to evaporate the HFIP. To the resulting thin film of peptide, dimethylsulfoxide (DMSO; Aldrich Chemical Co., Inc.) (25 μl) was added to dissolve the peptide. A 2 mm×7 mm TEFLON™-coated magnetic stir bar was added to each tube. Buffer (475 μL of 100 mM NaCl, 10 mM sodium phosphate, pH 7.4) was added to the DMSO solution with stirring. The resulting mixture was stirred continuously and the optical density was monitored at 400 nm to observe the formation of insoluble peptide aggregates.

Alternatively, β-AP_{1-40} was dissolved in DMSO as described above at 1.6 mM (6.9 mg/ml) and aliquots (25 μl) were added to stirred buffer (475 μl), followed by monitoring of absorbance at 400 nm.

For inhibition studies in which a β-amyloid modulator was dissolved in solution together with the natural β-AP, the modulators were dissolved in DMSO either with or without prior dissolution in HFIP. These compounds were then added to buffer with stirring, followed by addition of βAP_{1-40} in DMSO. Alternatively, HFIP solutions of modulators were combined with β-AP_{1-40} in HFIP followed by evaporation and redissolution of the mixture in DMSO. Buffer was then added to the DMSO solution to initiate the assay. The amino-terminally biotinylated β-amyloid peptide modulators N-biotinyl-βAP_{1-40}, and N-biotinyl-βAP_{1-15} were tested at concentrations of 1% and 5% in the natural β-AP_{1-40} solution.

A representative example of the results is shown graphically in FIG. 1, which depicts the inhibition of aggregation of natural β-AP_{1-40} by N-biotinyl-βAP_{1-40}. In the absence of the modulator, the optical density of the natural β-AP solution showed a characteristic sigmoidal curve, with a lag time prior to aggregation (approximately 3 hours in FIG. 1) in which the $A_{400\ nm}$ was low, followed by rapid increase in the $A_{400\ nm}$, which quickly reached a plateau level, representing aggregation of the natural β amyloid peptides. In contrast, in the presence of as little as 1% of the N-biotinyl-βAP_{1-40} modulator, aggregation of the natural β amyloid peptides was markedly inhibited, indicated by an increase in the lag time, a decrease in the slope of aggregation and a decrease in the plateau level reached for the turbidity of the solution (see FIG. 1). N-biotinyl-βAP_{1-40} at a concentration of 5% similarly inhibited aggregation of the natural β amyloid peptide. Furthermore, similar results were observed when N-biotinyl-βAP_{1-15} was used as the modulator. These results demonstrate that an N-terminally biotinylated β-AP modulator can effectively inhibit the aggregation of natural β amyloid peptides, even when the natural β amyloid peptides are present at as much as a 100-fold molar excess concentration.

B. Fluorescence Assay

In a second assay, β-AP aggregation was measured using a fluorometric assay essentially as described in Levine, H. (1993) *Protein Science* 2:404–410. In this assay, the dye thioflavine T (ThT) is contacted with the β-AP solution. Association of ThT with aggregated β-AP, but not monomeric or loosely associated β-AP, gives rise to a new excitation (ex) maximum at 450 nm and an enhanced emission (em) at 482 nm, compared to the 385 nm (ex) and 445 nm (em) for the free dye. β-AP aggregation was assayed by this method as follows. Aliquots (2.9 μl) of the solutions used in the aggregation assays as described above in section A were removed from the samples and diluted in 200 μl of potassium phosphate buffer (50 mM, pH 7.0) containing thioflavin T (10 μM; obtained commercially from Aldrich Chemical Co., Inc.). Excitation was set at 450 nm and emission was measured at 482 nm. Similar to the results observed with the optical density assay described above in section A, as little as 1% of the N-biotinylated β-AP modulators was effective at inhibiting the aggregation of natural β amyloid peptides using this fluorometric assay.

C. Static Aggregation Assay

In a third assay, β-AP aggregation was measured by visualization of the peptide aggregates using SDS-polyacrylamide gel electrophoresis (SDS-PAGE). In this assay, β-AP solutions were allowed to aggregate over a period of time and then aliquots of the reaction were run on a standard SDS-PAGE gel. Typical solution conditions were 200 μM of β-AP_{1-40} in PBS at 37° C. for 8 days or 200 μM β-AP_{1-40} in 0.1M sodium acetate at 37° C. for 3 days. The peptide aggregates were visualized by Coomassie blue staining of the gel or, for β-AP solutions that included a biotinylated β-AP modulator, by western blotting of a filter prepared from the gel with a streptavidin-peroxidase probe, followed by a standard peroxidase assay. The β-AP aggregates are identifiable as high molecular weight, low mobility bands on the gel, which are readily distinguishable from the low molecular weight, high mobility β-AP monomer or dimer bands.

When natural β-AP_{1-40} aggregation was assayed by this method in the absence of any β amyloid modulators, high molecular weight aggregates were readily detectable on the gel. In contrast, when N-biotinyl-β-AP_{1-40} modulator self-aggregation was assayed (i.e., aggregation of the N-biotinyl peptide alone, in the absence of any natural β-AP), few if any high molecular weight aggregates were observed, indicating that the ability of the modulator to self-aggregate is significantly reduced compared to natural β-AP. Finally, when aggregation of a mixture of natural β-AP_{1-40} and N-biotinylated β-AP_{1-40} was assayed by this method, reduced amounts of the peptide mixture associated into high molecular weight aggregates, thus demonstrating that the β amyloid modulator is effective at inhibiting the aggregation of the natural D amyloid peptides.

EXAMPLE 3

Neurotoxicity Analysis of β-Amyloid Modulators

The neurotoxicity of the β-amyloid modulators is tested in a cell-based assay using the neuronal precursor cell line PC-12, or primary neuronal cells, and the viability indicator 3,(4,4-dimethylthiazol-2-yl)2,5-diphenyl-tetrazolium bromide (MTT). (See Shearman, M. S. et al. (1994) *Proc. Natl. Acad. Sci. USA* 91:1470–1474; Hansen, M. B. et al. (1989) *J. Immun. Methods* 119:203–210). PC-12 is a rat adrenal pheochromocytoma cell line and is available from the American Type Culture Collection, Rockville, Md. (ATCC CRL 1721). MTT (commercially available from Sigma Chemical Co.) is a chromogenic substrate that is converted from yellow to blue in viable cells, which can be detected spectrophotometrically.

To test the neurotoxicity of a β-amyloid modulator (either alone or combined with natural β-AP), cells first are plated in 96-well plates at 7,000–10,000 cells/well and allowed to adhere by overnight culture at 37° C. Serial dilutions of freshly dissolved or "aged" modulators (either alone or combined with natural β-AP) in phosphate buffered saline (PBS) are added to the wells in triplicate and incubation is continued for two or more days. Aged modulators are prepared by incubating an aqueous solution of the modulator at 37° C. undisturbed for a prolonged period (e.g., five days or more). For the final two hours of exposure of the cells to the modulator preparation, MTT is added to the media to a final concentration of 1 mg/ml and incubation is continued at 37° C. Following the two hour incubation with MTT, the media is removed and the cells are lysed in isopropanol/0.4N HCl with agitation. An equal volume of PBS is added to each well and the absorbance of each well at 570 nm is measured to quantitate viable cells. Alternatively, MTT is solubilized by addition of 50% N,N-dimethyl formamide/20% sodium dodecyl sulfate added directly to the media in the wells and viable cells are likewise quantitated by measuring absorbance at 570 nm. The relative neurotoxicity of a β-amyloid modulator (either alone or in combination with natural β-AP) is determined by comparison to natural β-AP alone (e.g., β1–40, β1–42), which exhibits neurotoxicity in this assay and thus can serve as a positive control.

EXAMPLE 4

Synthesis of Additional Modified β-Amyloid Peptide Compounds

In this example, a series of modified β-APs, having a variety of N-terminal or random side chain modifications were synthesized.

A series of N-terminally modified β-amyloid peptides was synthesized using standard methods. Fully-protected resin-bound peptides corresponding to Aβ(1–15) and Aβ(1–40) were prepared as described in Example 1 on Wang resin to eventually afford carboxyl terminal peptide acids. Small portions of each peptide resin (13 and 20 μmoles, respectively) were aliquoted into the wells of the reaction block of an Advanced ChemTech Model 396 Multiple Peptide Synthesizer. The N-terminal FMOC protecting group of each sample was removed in the standard manner with 25% piperidine in NMP followed by extensive washing with NMP. The unprotected N-terminal α-amino group of each peptide-resin sample was modified using one of the following methods:

Method A, coupling of modifying reagents containing free carboxylic acid groups: The modifying reagent (five equivalents) was predissolved in NMP, DMSO or a mixture of these two solvents. HOBT and DIC (five equivalents of each reagent) were added to the dissolved modifier and the resulting solution was added to one equivalent of free-amino peptide-resin. Coupling was allowed to proceed overnight, followed by washing. If a ninhydrin test on a small sample of peptide-resin showed that coupling was not complete, the coupling was repeated using 1-hydroxy-7-azabenzotriazole (HOAt) in place of HOBt.

Method B, coupling of modifying reagents obtained in preactivated forms: The modifying reagent (five equivalents) was predissolved in NMP, DMSO or a mixture of these two solvents and added to one equivalent of peptide-resin. Diisopropylethylamine (DIEA; six equivalents) was added to the suspension of activated modifier and peptide-resin. Coupling was allowed to proceed overnight, followed by washing. If a ninhydrin test on a small sample of peptide-resin showed that coupling was not complete, the coupling was repeated.

After the second coupling (if required) the N-terminally modified peptide-resins were dried at reduced pressure and cleaved from the resin with removal of side-chain protecting groups as described in Example 1. Analytical reversed-phase HPLC was used to confirm that a major product was present in the resulting crude peptides which were purified using Millipore Sep-Pak cartridges or preparative reverse-phase HPLC. Mass spectrometry was used to confirm the presence of the desired compound in the product.

Method A was used to couple N-acetylneuraminic acid, cholic acid, trans-4-cotininecarboxylic acid, 2-imino-1-imidazolidineacetic acid, (S)-(−)-indoline-2-carboxylic acid, (−)-menthoxyacetic acid, 2-norbornaneacetic acid, γ-oxo-5-acenaphthenebutyric acid, (−)-2-oxo-4-thiazolidinecarboxylic acid, and tetrahydro-3-furoic acid. Method B was used to couple 2-iminobiotin-N-hydroxysuccinimide ester, diethylenetriaminepentaacetic dianhydride, 4-morpholinecarbonyl chloride, 2-thiopheneacetyl chloride, and 2-thiophenesulfonyl chloride.

In a manner similar to the construction of N-terminally modified Aβ(1–15) and Aβ(1–40) peptides described above, N-fluoresceinyl Aβ(1–15) and Aβ(1–40) were prepared in two alternative manners using the preactivated reagents 5-(and 6)-carboxyfluorescein succinimidyl ester and fluorescein-5-isothiocyanate (FITC Isomer I). Both reagents were obtained from Molecular Probes Inc. Couplings were performed using four equivalents of reagent per equivalent of peptide-resin with DIEA added to make the reaction solution basic to wet pH paper. Couplings of each reagent to Aβ(1–15)-resin appeared to be complete after a single

overnight coupling. Coupling to Aβ(1–40)-resin was slower as indicated by a positive ninhydrin test and both reagents were recoupled to this peptide-resin overnight in tetrahydrofuran-NMP (1:2 v/v). The resulting N-terminally modified peptide-resins were cleaved, deprotected and purified as described in Example A.

In addition to the N-fluoresceinyl Aβ peptides described above, a β-amyloid modulator comprised of random modification of Aβ(1–40) with fluorescein was prepared. Aβ(1–40) purchased from Bachem was dissolved in DMSO at approximately 2 mg/mL. 5-(and-6)-Carboxyfluorescein purchased from Molecular Probes was added in a 1.5 molar excess and DIEA was added to make the solution basic to wet pH paper. The reaction was allowed to proceed for 1 hour at room temperature and was then quenched with triethanolamine. The product was added to assays as this crude mixture.

β-amyloid modulator compounds comprising other regions of the β-AP amino acid sequence (e.g., an Aβ aggregation core domain) were similarly prepared using the synthesis methods described above. Moreover, modulators comprising other amyloidogenic peptides can be similarly prepared.

EXAMPLE 5

Identification of Additional β-Amyloid Modulators

In this Example, two assays of Aβ aggregation were used to identify β-amyloid modulators which can inhibit this process.

The first assay is referred to as a seeded static assay (SSA) and was performed as follows:

To prepare a solution of Aβ monomer, the appropriate quantity of Aβ(1–40) peptide (Bachem) was weighed out on a micro-balance (the amount was corrected for the amount of water in the preparation, which, depending on lot number, was 20–30% w/w). The peptide was dissolved in ½5 volume of dimethysulfoxide (DMSO), followed by water to ½ volume and ½ volume 2× PBS (10× PBS: NaCl 137 mM, KCl 2.7 mM $Na_2HPO_4 \cdot 7H_2O$ 4.3 mM, KH_2PO_4 1.4 mM pH 7.2) to a final concentration of 200 μM.

To prepare a stock seed, 1 ml of the above Aβ monomer preparation, was incubated for 8 days at 37° C. and sheared sequentially through an 18, 23, 26 and 30 gauge needle 25, 25, 50, and 100 times respectively. 2 μl samples of the sheared material was taken for fluorescence measurements after every 50 passes through the 30 gauge needle until the fluorescence units (FU) had plateaued (approx. 100–150×).

To prepare a candidate inhibitor, the required amount of candidate inhibitor was weighed out and the stock dissolved in 1× PBS to a final concentration of 1 mM (10× stock). If insoluble, it was dissolved in ⅒ volume of DMSO and diluted in 1× PBS to 1 mM. A further ⅒ dilution was also prepared to test each candidate at both 100 μM and 10 μM.

For the aggregation assay, each sample was set up in triplicate [50 μl of 200 μM monomer, 125 FU sheared seed (variable quantity dependent on the batch of seed, routinely 3–6 μl), 10 μl of 10× inhibitor solution, final volume made up to 100 μl with 1× PBS]. Two concentrations of each inhibitor were tested 100 μM and 10 μM, equivalent to a 1:1 and a 1:10 molar ratio of monomer to inhibitor. The controls included an unseeded reaction to confirm that the fresh monomer contained no seed, and a seeded reaction in the absence of inhibitor, as a reference to compare against putative inhibitors. The assay was incubated at 37° C. for 6 h, taking 2 μl samples hourly for fluorescence measurements. To measure fluorescence, a 2 μl sample of Aβ was added to 400 μl of Thioflavin-T solution (50 mM Potassium Phosphate 10 mM Thioflavin-T pH 7.5). The samples were vortexed and the fluorescence was read in a 0.5 ml micro quartz cuvette at EX 450 nm and EM 482 nm (Hitachi 4500 Fluorimeter). β-aggregation results in enhanced emission of Thioflavin-T. Accordingly, samples including an effective inhibitor compound exhibit reduced emission as compared to control samples without the inhibitor compound.

The second assay is referred to as a shaken plate aggregation assay and was performed as follows:

Aβ(1–40) peptide from Bachem (Torrance, Calif.) was dissolved in HFIP (1,1,1,3,3,3-Hexafluoro-2-propanol; Aldrich 10,522-8) at a concentration of 2 mg peptide/ml and incubated at room temperature for 30 min. HFIP solubilized peptide was sonicated in a waterbath sonicator for 5 min at highest setting, then evaporated to dryness under a stream of argon. The peptide film was resuspended in anhydrous dimethylsulfoxide (DMSO) at a concentration of 6.9 mg/ml, sonicated for 5 min as before, then filtered through a 0.2 micron nylon syringe filter (VWR cat. No. 28196-050). Candidate inhibitors were dissolved directly in DMSO, generally at a molar concentration 4 times that of the Aβ(1–40) peptide.

Candidates were assayed in triplicate. For each candidate to be tested, 4 parts Aβ(1–40) peptide in DMSO were combined with 1 part candidate inhibitor in DMSO in a glass vial, and mixed to produce a 1:1 molar ratio of Aβ peptide to candidate. For different molar ratios, candidates were diluted with DMSO prior to addition to Aβ(1-40), in order to keep the final DMSO and Aβ(1–40) concentrations constant. Into an ultra low binding 96 well plate (Corning Costar cat. No. 2500, Cambridge Mass.) 100 μl PTL buffer (150 mM NaCl, 10 mM NaH_2PO_4; pH 7.4) was aliquotted per well. For each candidate, 10 μl of peptide mixture in DMSO was aliquotted into each of three wells containing buffer. The covered plate was vigorously vortexed on a plate shaker at high speed for 30 seconds. An additional 100 μl of PTL buffer was added to each well and again the plate was vortexed vigorously for 30 sec. Absorbance at 405 nm was immediately read in a plate reader for a baseline reading. The plate was returned to the plate shaker and vortexed at moderate speed for 5 hours at room temperature, with absorbance readings taken at 15–20 min intervals. Increased absorbance indicated aggregation. Accordingly, effective inhibitor compounds cause a decrease in absorbance in the test sample as compared to a control sample without the inhibitor compound.

Representative results of the static seeded assay and shaken plate assay with preferred β-amyloid modulators are shown below in Table I.

TABLE I

Candidate Inhibitor	Aβ Amino Acids	Modifying Reagent	Effect in shaken plate assay	Effect in Seeded Static Assay*
174	Aβ1-15	Cholic acid	Complete inhibition at 100% conc	++
176	Aβ1-15	Diethylene-triamine penta acetic acid	Decreased Plateau	++
180	Aβ1-15	(−)-Menthoxy	None	++

TABLE I-continued

Candidate Inhibitor	Aβ Amino Acids	Modifying Reagent	Effect in shaken plate assay	Effect in Seeded Static Assay*
		acetic acid		
190	Aβ1-15	Fluorescein carboxylic acid (FICO)	Decreased Plateau	++
220	Aβ16-40 mutant	NH_2—EVHHHHQQK-[Aβ(16–40)]-COOH (SEQ ID NO:16)	Complete inhibition at 100%, increased lag at 10%	++
224	Aβ1-40 mutant	$F_{19}F_{20}$ -> $T_{19}T_{20}$	Increased lag	++
233	A6β-20	Acetic acid	accelerated aggregation at 10% conc	++

*++ = A strong inhibitor of aggregation. The rate of aggregation in the presence of the inhibitor was decreased compared to the control by at least 30 . 50%

These results indicate that β-APs modified by a wide variety of N-terminal modifying groups are effective at modulating β-amyloid aggregation.

EXAMPLE 6

Additional β-Amyloid Aggregation Assays

Most preferably, the ability of β-amyloid modulator compounds to modulate (e.g., inhibit or promote) the aggregation of natural β-AP when combined with the natural β-AP is examined in one or both of the aggregation assays described below. Natural β-AP (β-AP_{1-40}) for use in the aggregation assays is commercially available from Bachem (Torrance, Calif.).

A. Nucleation Assay

The nucleation assay is employed to determine the ability of test compounds to alter (e.g. inhibit) the early events in formation of β-AP fibers from monomeric β-AP. Characteristic of a nucleated polymerization mechanism, a lag time is observed prior to nucleation, after which the peptide rapidly forms fibers as reflected in a linear rise in turbidity. The time delay before polymerization of β-AP monomer can be quantified as well as the extent of formation of insoluble fiber by light scattering (turbidity). Polymerization reaches equilibrium when the maximum turbidity reaches a plateau. The turbidity of a solution of natural β-AP in the absence or presence of various concentrations of a β-amyloid modulator compound is determined by measuring the apparent absorbance of the solution at 405 nm ($A_{405\ nm}$) over time. The threshold of sensitivity for the measurement of turbidity is in the range of 15–20 μM β-AP. A decrease in turbidity over time in the presence of the modulator, as compared to the turbidity in the absence of the modulator, indicates that the modulator inhibits formation of β-AP fibers from monomeric β-AP. This assay can be performed using stirring or shaking to accelerate polymerization, thereby increasing the speed of the assay. Moreover the assay can be adapted to a 96-well plate format to screen multiple compounds.

To perform the nucleation assay, first $Aβ_{1-40}$ peptide is dissolved in HFIP (1,1,1,3,3,3-Hexafluoro-2-propanol; Aldrich 10,522-8) at a concentration of 2 mg peptide/ml and incubated at room temperature for 30 min. HFIP-solubilized peptide is sonicated in a waterbath sonicator for 5 min at highest setting, then evaporated to dryness under a stream of argon. The peptide film is resuspended in anhydrous dimethylsulfoxide (DMSO) at a concentration of 6.9 mg/ml (25× concentration), sonicated for 5 min as before, then filtered through a 0.2 micron nylon syringe filter (VWR cat. No. 28196-050). Test compounds are dissolved in DMSO at a 100× concentration. Four volumes of 25× $Aβ_{1-40}$ peptide in DMSO are combined with one volume of test compound in DMSO in a glass vial, and mixed to produce a 1:1 molar ratio of Aβ peptide to test compound. For different molar ratios, test compounds are diluted with DMSO prior to addition to $Aβ_{1-40}$, in order to keep the final DMSO and $Aβ_{1-40}$ concentrations constant. Control samples do not contain the test compound. Ten microliters of the mixture is then added to the bottom of a well of a Corning Costar ultra low binding 96-well plate (Corning Costar, Cambridge Mass.; cat. No. 2500). Ninety microliters of water is added to the well, the plate is shaken on a rotary shaken at a constant speed at room temperature for 30 seconds, an additional 100 μl of 2× PTL buffer (20 mM NaH_2PO_4, 300 mM NaCl, pH 7.4) is added to the well, the plate is reshaken for 30 seconds and a baseline (t=0) turbidity reading is taken by measuring the apparent absorbance at 405 nm using a Bio-Rad Model 450 Microplate Reader. The plate is then returned to the shaker and shaken continuously for 5 hours. Turbidity readings are taken at 15 minute intervals.

β-amyloid aggregation in the absence of any modulators results in enhanced turbidity of the natural β-AP solution (i.e., an increase in the apparent absorbance at 405 nm over time). Accordingly, a solution including an effective inhibitory modulator compound exhibits reduced turbidity as compared to the control sample without the modulator compound (i.e., less apparent absorbance at 405 nm over time as compared to the control sample).

B. Seeded Extension Assay

The seeded extension assay can be employed to measure the rate of Aβ fiber formed in a solution of Aβ monomer following addition of polymeric Aβ fiber "seed". The ability of test compounds to prevent further deposition of monomeric Aβ to previously deposited amyloid is determined using a direct indicator of β-sheet formation using fluorescence. In contrast with the nucleation assay, the addition of seed provides immediate nucleation and continued growth of preformed fibrils without the need for continuous mixing, and thus results in the absence of a lag time before polymerization starts. Since this assay uses static polymerization conditions, the activity of positive compounds in the nucleation assay can be confirmed in this second assay under different conditions and with an additional probe of amyloid structure.

In the seeded extension assay, monomeric $Aβ_{1-40}$ is incubated in the presence of a "seed" nucleus (approximately ten mole percent of Aβ that has been previously allowed to polymerize under controlled static conditions). Samples of the solution are then diluted in thioflavin T (Th-T). The polymer-specific association of Th-T with Aβ produces a fluorescent complex that allows the measurement of the extent of fibril formation (Levine, H. (1993) *Protein Science* 2:404–410). In particular, association of Th-T with aggregated β-AP, but not monomeric or loosely associated β-AP, gives rise to a new excitation (ex) maximum at 450 nm and an enhanced emission (em) at 482 nm, compared to the 385 nm (ex) and 445 nm (em) for the free dye. Small aliquots of the polymerization mixture contain sufficient fibril to be mixed with Th-T to allow the monitoring of the reaction mixture by repeated sampling. A linear growth curve is observed in the presence of excess monomer. The formation of thioflavin T responsive β-sheet

fibrils parallels the increase in turbidity observed using the nucleation assay.

A solution of Aβ monomer for use in the seeded extension assay is prepared by dissolving an appropriate quantity of $A\beta_{1-40}$ peptide in ⅛s volume of dimethylsulfoxide (DMSO), followed by water to ½ volume and ½ volume 2× PBS (10× PBS: NaCl 137 mM, KCl 2.7 mM $Na_2HPO_4.7H_2O$ 4.3 mM, KH_2PO_4 1.4 mM pH 7.2) to a final concentration of 200 μM. To prepare the stock seed, 1 ml of the Aβ monomer preparation, is incubated for approximately 8 days at 37° C. and sheared sequentially through an 18, 23, 26 and 30 gauge needle 25, 25, 50, and 100 times respectively. 2 μl samples of the sheared material is taken for fluorescence measurements after every 50 passes through the 30 gauge needle until the fluorescence units (FU) plateau (approx. 100–150×). Test compounds are prepared by dissolving an appropriate amount of test compound in 1× PBS to a final concentration of 1 mM (10× stock). If insoluble, the compound is dissolved in 1/10 volume of DMSO and diluted in 1× PBS to 1 mM. A further 1/10 dilution is also prepared to test each candidate at both 100 μM and 10 μM.

To perform the seeded extension assay, each sample is set up with 50 μl of 200 μM monomer, 125 FU sheared seed (a variable quantity dependent on the batch of seed, routinely 3–6 μl) and 10 μl of 10× modulator solution. The sample volume is then adjusted to a final volume of 100 μl with 1× PBS. Two concentrations of each modulator typically are tested: 100 μM and 10 μM, equivalent to a 1:1 and a 1:10 molar ratio of monomer to modulator. The controls include an unseeded reaction to confirm that the fresh monomer contains no seed, and a seeded reaction in the absence of any modulators, as a reference to compare against candidate modulators. The assay is incubated at 37° C. for 6 h, taking 2 μl samples hourly for fluorescence measurements. To measure fluorescence, a 2 μl sample of Aβ is added to 400 μl of Thioflavin-T solution (50 mM Potassium Phosphate 10 mM Thioflavin-T pH 7.5). The samples are vortexed and the fluorescence is read in a 0.5 ml micro quartz cuvette at EX 450 nm and EM 482 nm (Hitachi 4500 Fluorimeter).

β-amyloid aggregation results in enhanced emission of Thioflavin-T. Accordingly, samples including an effective inhibitory modulator compound exhibit reduced emission as compared to control samples without the modulator compound.

EXAMPLE 7

Effect of Different Amino Acid Subregions of Aβ Peptide on the Inhibitory Activity of β-Amyloid Modulator Compounds

To determine the effect of various subregions of $A\beta_{1-40}$ on the inhibitory activity of a a β-amyloid modulator, overlapping Aβ peptide 15 mers were constructed. For each 15 mer, four different amino-terminal modifiers were tested: a cholyl group, an iminobiotinyl group, an N-acetyl neuraminyl group (NANA) and a 5-(and 6-)-carboxyfluoresceinyl group (FICO). The modulators were evaluated in the nucleation and seeded extension assays described in Example 6.

The results of the nucleation assays are summarized below in Table II. The concentration of $A\beta_{1-40}$ used in the assays was 50 μM. The "mole %" value listed in Table II refers to the % concentration of the test compound relative to $A\beta_{1-40}$. Accordingly, 100% indicates that $A\beta_{1-40}$ and the test compound were equimolar. Mole % values less than 100% indicate that $A\beta_{1-40}$ was in molar excess relative to the test compound (e.g., 10% indicates that $A\beta_{1-40}$ was in 10-fold molar excess relative to the test compound). The results of the nucleation assays for each test compound are presented in Table II in two ways. The "fold increase in lag time", which is a measure of the ability of the compound to delay the onset of aggregation, refers to the ratio of the observed lag time in the presence of the test compound to the observed lag time in the control without the test compound. Accordingly a fold increase in lag time of 1.0 indicates no change in lag time, whereas numbers >1.0 indicate an increase in lag time. The "% inhibition of plateau", which is a measure of the ability of the compound to decrease the total amount of aggregation, refers to the reduction of the final turbidity in the presence of the test compound expressed as a percent of the control without the test compound. Accordingly, an inhibitor that abolishes aggregation during the course of the assay will have a % inhibition of 100. N-terminally modified Aβ subregions which exhibited inhibitory activity are indicated in bold in Table II.

TABLE II

Reference #	N-terminal Modification	Aβ Peptide	Mole %	Fold Increase in Lag Time	% Inhibition of Plateau
PPI-174	cholyl	$A\beta_{1-15}$	100	>4.5	100
PPI-264	cholyl	$A\beta_{6-20}$	100	>4.5	100
PPI-269	cholyl	$A\beta_{11-25}$	100	1.5	~0
PPI-274	cholyl	$A\beta_{16-30}$	100	>4.5	100
PPI-279	cholyl	$A\beta_{21-35}$	100	1.6	51
PPI-284	cholyl	$A\beta_{26-40}$	100	>4.5	87
PPI-173	NANA	$A\beta_{1-15}$	100	~1	~0
PPI-266	NANA	$A\beta_{6-20}$	100	1.3	64
PPI-271	NANA	$A\beta_{11-25}$	100	1.3	77
PPI-276	NANA	$A\beta_{16-30}$	100	~1	~0
PPI-281	NANA	$A\beta_{21-35}$	100	~1	53
PPI-286	NANA	$A\beta_{26-40}$	100	1.3	~0
PPI-172	Iminobiotinyl	$A\beta_{1-15}$	100	1.2	~0
PPI-267	Iminobiotinyl	$A\beta_{6-20}$	100	1.6	44
PPI-272	Iminobiotinyl	$A\beta_{11-25}$	100	1.2	40
PPI-277	Iminobiotinyl	$A\beta_{16-30}$	100	1.2	55
PPI-282	Iminobiotinyl	$A\beta_{21-35}$	100	~1	66
PPI-287	Iminobiotinyl	$A\beta_{26-40}$	100	2.3	~0
PPI-190	FICO	$A\beta_{1-15}$	100	~1	30
PPI-268	FICO	$A\beta_{6-20}$	100	1.9	~0
PPI-273	FICO	$A\beta_{11-25}$	100	1.7	34
PPI-278	FICO	$A\beta_{16-30}$	100	1.6	59
PPI-283	FICO	$A\beta_{21-35}$	100	1.2	25
PPI-288	FICO	$A\beta_{26-40}$	100	2	75

These results indicate that certain subregions of $A\beta_{1-40}$, when modified with an appropriate modifying group, are effective at inhibiting the aggregation of $A\beta_{1-40}$. A cholyl group was an effective modifying group for several subregions. Cholic acid alone was tested for inhibitory activity but had no effect on Aβ aggregation. The $A\beta_{6-20}$ subregion exhibited high levels of inhibitory activity when modified with several different modifying groups (cholyl, NANA, iminobiotinyl), with cholyl-$A\beta_{6-20}$ (PPI-264) being the most active form. Accordingly, this modulator compound was chosen for further analysis, described in Example 8.

EXAMPLE 8

Identification of a Five Amino Acid Subregion of Aβ Peptide Sufficient for Inhibitory Activity of a β-Amyloid Modulator Compound

To further delineate a minimal subregion of cholyl-$A\beta_{6-20}$ sufficient for inhibitory activity, a series of amino terminal and carboxy terminal amino acid deletions of cholyl-$A\beta_{6-20}$ were constructed. The modulators all had the same cholyl amino-terminal modification. Additionally, for the peptide series having carboxy terminal deletions, the carboxy ter-

minus was further modified to an amide. The modulators were evaluated as described in Example 7 and the results are summarized below in Table III, wherein the data is presented as described in Example 7.

TABLE III

Ref. #	N-Term. Mod.	Aβ Peptide	C-Term. Mod.	Mole %	Fold Increase in Lag Time	% Inhibition of Plateau
PPI-264	cholyl	$Aβ_{6-20}$	—	100	>4.5	100
				10	2	43
PPI-341	cholyl	$Aβ_{7-20}$	—	100	>4.5	100
				33	2	~0
PPI-342	cholyl	$Aβ_{8-20}$	—	100	1.5	~0
				33	2.1	~0
PPI-343	cholyl	$Aβ_{9-20}$	—	33	2.0	~0
PPI-344	cholyl	$Aβ_{10-20}$	—	33	2.1	~0
PPI-345	cholyl	$Aβ_{11-20}$	—	33	1.5	~0
PPI-346	cholyl	$Aβ_{12-20}$	—	33	2.1	~0
PPI-347	cholyl	$Aβ_{13-20}$	—	33	2.6	~0
PPI-348	cholyl	$Aβ_{14-20}$	—	33	2.0	49
PPI-349	cholyl	$Aβ_{15-20}$	—	33	2.3	50
PPI-350	cholyl	$Aβ_{16-20}$	—	38	3.4	23
PPI-296	cholyl	$Aβ_{6-20}$	amide	33	1.8	~0
PPI-321	cholyl	$Aβ_{6-19}$	amide	33	1.4	~0
PPI-325	cholyl	$Aβ_{6-17}$	amide	33	1.8	~0
PPI-331	cholyl	$Aβ_{6-14}$	amide	33	1.0	29
PPI-339	cholyl	$Aβ_{6-10}$	amide	33	1.1	13

These results indicate that activity of the modulator is maintained when amino acid residue 6 is removed from the amino terminal end of the modulator (i.e., cholyl-$Aβ_{7-20}$ retained activity) but activity is lost as the peptide is deleted further at the amino-terminal end by removal of amino acid position 7 through to amino acid position 12 (i.e., cholyl-$Aβ_{8-20}$ through cholyl-$Aβ_{13-20}$ did inhibit the plateau level of Aβ aggregation). However, further deletion of amino acid position 13 resulted in a compound (i.e., cholyl-$Aβ_{4-20}$) in which inhibitory activity is restored. Furthermore, additional deletion of amino acid position 14 (i.e., cholyl-$Aβ_{15-20}$) or positions 14 and 15 (i.e., cholyl-$Aβ_{16-20}$) still maintained inhibitory activity. Thus, amino terminal deletions of $Aβ_{6-20}$ identified $Aβ_{16-20}$ as a minimal subregion which is sufficient for inhibitory activity when appropriately modified. In contrast, carboxy terminal deletion of amino acid position 20 resulted in loss of activity which was not fully restored as the peptide was deleted further at the carboxy-terminal end. Thus, maintenance of position 20 within the modulator may be important for inhibitory activity.

EXAMPLE 9

Identification of a Four Amino Acid Subregion of Aβ Peptide Sufficient for Inhibitory Activity of a β-Amyloid Modulator Compound

In this example, the smallest effective modulator identified in the studies described in Example 8, cholyl-$Aβ_{16-20}$ (PPI-350), was analyzed further. Additional amino- and carboxy-terminal deletions were made with cholyl-$Aβ_{16-20}$, as well as an amino acid substitution (Val_{18}->Thr), to identify the smallest region sufficient for the inhibitory activity of the modulator. A peptide comprised of five alanine residues, $(Ala)_5$; SEQ ID NO:35, modified at its amino-terminus with cholic acid, was used as a specificity control. The modulators were evaluated as described in Example 7 and the results are summarized below in Table IV, wherein the data is presented as described in Example 7.

TABLE IV

Ref. #	N-Term. Mod.	Aβ Peptide	C-Term. Mod.	Mole %	Fold Increase in Lag Time	% Inhibition of Plateau
PPI-264	cholyl	$Aβ_{6-20}$	—	10	2.0	43
PPI-347	cholyl	$Aβ_{13-20}$	—	10	2.2	57
PPI-349	cholyl	$Aβ_{15-20}$	—	100	>5.0	100
				33	2.6	35
				10	2.1	~0
PPI-350	cholyl	$Aβ_{16-20}$	—	100	>5.0	100
				10	2.4	40
PPI-368	cholyl	$Aβ_{17-21}$	—	100	>5.0	100
PPI-374	imino-biotinyl	$Aβ_{16-20}$	—	100	1.3	86
PPI-366	cholyl	$Aβ_{15-19}$	—	100	3.1	~0
				10	1.6	~0
PPI-369	cholyl	$Aβ_{16-20}$ (Val_{18}-> Thr)	—	100	~1	~0
PPI-370	cholyl	$Aβ_{16-20}$ (Phe_{19}-> Ala)	—	100	2.6	73
PPI-365	cholyl	$(Ala)_5$	—	100	~1	~0
PPI-319	cholyl	$Aβ_{16-20}$	amide	33	5.6	~0
				10	2.7	~0
PPI-321	cholyl	$Aβ_{16-19}$	amide	100	1.2	~0
PPI-377	—	$Aβ_{16-20}$	—	100	~1	~0

As shown in Table IV, cholyl-$Aβ_{16-20}$ (PPI-350) and cholyl-$Aβ_{17-21}$ (PPI-368) both exhibited inhibitory activity, indicating that the four-amino acid minimal subregion of positions 17–20 is sufficient for inhibitory activity. Loss of position 20 (e.g., in PPI-366 and PPI-321) resulted in loss of inhibitory activity, demonstrating the importance of position 20. Moreover, mutation of valine at position 18 to threonine (in PPI-369) also resulted in loss of activity, demonstrating the importance of position 18. In contrast, mutation of phenylalanine at position 19 to alanine (cholyl-$Aβ_{16-20}$ Phe_{19}->Ala; PPI-370) resulted in a compound which still retained detectable inhibitory activity. Accordingly, the phenylalanine at position 19 is more amenable to substitution, preferably with another hydrophobic amino acid residue. Cholyl-penta-alanine; SEQ ID NO:35 (PPI-365) showed no inhibitory activity, demonstrating the specificity of the Aβ peptide portion of the modulator. Moreover, unmodified $Aβ_{16-20}$ (PPI-377) was not inhibitory, demonstrating the functional importance of the amino-terminal modifying group. The specific functional group influenced the activity of the modulator. For example, iminobiotinyl-$Aβ_{16-20}$ (PPI-374) exhibited inhibitory activity similar to cholyl-$Aβ_{16-20}$, whereas an N-acetyl neuraminic acid (NANA)-modified $Aβ_{16-20}$ was not an effective inhibitory modulator (not listed in Table IV). A C-terminal amide derivative of cholyl-$Aβ_{16-20}$ (PPI-319) retained high activity in delaying the lag time of aggregation, indicating that the carboxy-terminus of the modulator can be derivatized without loss of inhibitory activity. Although this amide-derivatized compound did not inhibit the overall plateau level of aggregation over time, the compound was not tested at concentrations higher than mole 33%. Higher concentrations of the amide-derivatized compound are predicted to inhibit the overall plateau level of aggregation, similar to cholyl-$Aβ_{16-20}$ (PPI-350).

EXAMPLE 10

Effect of β-Amyloid Modulators on the Neurotoxicity of Natural β-Amyloid Peptide Aggregates

The neurotoxicity of natural β-amyloid peptide aggregates, in either the presence or absence of a β-amyloid

modulator, is tested in a cell-based assay using either a rat or human neuronally-derived cell line (PC-12 cells or NT-2 cells, respectively) and the viability indicator 3,(4,4-dimethylthiazol-2-yl)2,5-diphenyl-tetrazolium bromide (MTT). (See e.g., Shearman, M. S. et al. (1994) *Proc. Natl. Acad. Sci. USA* 91:1470–1474; Hansen, M. B. et al. (1989) *J. Immun. Methods* 119:203–210 for a description of similar cell-based viability assays). PC-12 is a rat adrenal pheochromocytoma cell line and is available from the American Type Culture Collection, Rockville, Md. (ATCC CRL 1721). MTT (commercially available from Sigma Chemical Co.) is a chromogenic substrate that is converted from yellow to blue in viable cells, which can be detected spectrophotometrically.

To test the neurotoxicity of natural β-amyloid peptides, stock solutions of fresh Aβ monomers and aged Aβ aggregates were first prepared. $A\beta_{1-40}$ in 100% DMSO was prepared from lyophilized powder and immediately diluted in one half the final volume in H_2O and then one half the final volume in 2× PBS so that a final concentration of 200 μM peptide, 4% DMSO is achieved. Peptide prepared in this way and tested immediately on cells is referred to as "fresh" Aβ monomer. To prepare "aged" Aβ aggregates, peptide solution was placed in a 1.5 ml Eppendorf tube and incubated at 37° C. for eight days to allow fibrils to form. Such "aged" Aβ peptide can be tested directly on cells or frozen at −80° C. The neurotoxicity of fresh monomers and aged aggregates were tested using PC12 and NT2 cells. PC12 cells were routinely cultured in Dulbeco's modified Eagle's medium (DMEM) containing 10% horse serum, 5% fetal calf serum, 4 mM glutamine, and 1% gentamycin. NT2 cells were routinely cultured in OPTI-MEM medium (GIBCO BRL CAT. #31985) supplemented with 10% fetal calf serum, 2 mM glutamine and 1% gentamycin. Cells were plated at 10–15,000 cells per well in 90 μl of fresh medium in a 96-well tissue culture plate 3–4 hours prior to treatment. The fresh or aged Aβ peptide solutions (10 μL) were then diluted 1:10 directly into tissue culture medium so that the final concentration was in the range of 1–10 μM peptide. Cells are incubated in the presence of peptide without a change in media for 48 hours at 37° C. For the final three hours of exposure of the cells to the β-AP preparation, MTT was added to the media to a final concentration of 1 mg/ml and incubation was continued at 37° C. Following the two hour incubation with MTT, the media was removed and the cells were lysed in 100 μL isopropanol/0.4N HCl with agitation. An equal volume of PBS was added to each well and the plates were agitated for an additional 10 minutes. Absorbance of each well at 570 nm was measured using a microtiter plate reader to quantitate viable cells.

The neurotoxicity of aged (5 day or 8 day) $A\beta_{1-40}$ aggregates alone, but not fresh $A\beta_{1-40}$ monomers alone, was confirmed in an experiment the results of which are shown in FIG. 3, which demonstrates that incubating the neuronal cells with increasing amounts of fresh $A\beta_{1-40}$ monomers was not significantly toxic to the cells whereas incubating the cells with increasing amounts of 5 day or 8 day $A\beta_{1-40}$ aggregates led to increasing amount of neurotoxicity. The EC50 for toxicity of aged $A\beta_{1-40}$ aggregates was 1–2 μM for both the PC12 cells and the NT2 cells.

To determine the effect of a β-amyloid modulator compound on the neurotoxicity of $A\beta_{1-40}$ aggregates, a modulator compound, cholyl-$A\beta_{6-20}$ (PPI-264), was preincubated with $A\beta_{1-40}$ monomers under standard nucleation assay conditions as described in Example 6 and at particular time intervals post-incubation, aliquots of the β-AP/modulator solution were removed and 1) the turbidity of the solution was assessed as a measure of aggregation and 2) the solution was applied to cultured neuronal cells for 48 hours at which time cell viability was assessed using MTT to determine the neurotoxicity of the solution. The results of the turbidity analysis are shown in FIG. 4, panels A, B and C. In panel A, $A\beta_{1-40}$ and cholyl-$A\beta_{6-20}$ were both present at 64 μM. In panel B, $A\beta_{1-40}$ was present at 30 μM and cholyl-$A\beta_{6-20}$ was present at 64 μM. In panel C, $A\beta_{1-40}$ was present at 10 μM and cholyl-$A\beta_{6-20}$ was present at 64 μM. These data show that an equimolar amount of cholyl-$A\beta_{6-20}$ is effective at inhibiting aggregation of $A\beta_{1-40}$ (see FIG. 4, panel A) and that as the concentration of $A\beta_{1-40}$ is reduced, the amount of detectable aggregation of the $A\beta_{1-40}$ monomer is correspondingly reduced (compare FIG. 4, panels B and C with panel A). The corresponding results of the neurotoxicity analysis are shown in FIG. 4, panels D, E, and F. These results demonstrate that the β-amyloid modulator compound not only inhibits aggregation of $A\beta_{1-40}$ monomers but also inhibits the neurotoxicity of the $A\beta_{1-40}$ solution, illustrated by the reduced percent toxicity of the cells when incubated with the $A\beta_{1-40}$/modulator solution as compared to $A\beta_{1-40}$ alone (see e.g., FIG. 4, panel D). Moreover, even when $A\beta_{1-40}$ aggregation was not detectable as measured by light scattering, the modulator compound inhibited the neurotoxicity of the $A\beta_{1-40}$ solution (see FIG. 4, panels E and F). Thus, the formation of neurotoxic $A\beta_{1-40}$ aggregates precedes the formation of insoluble aggregates detectable by light scattering and the modulator compound is effective at inhibiting the inhibiting the formation and/or activity of these neurotoxic aggregates. Similar results were seen with other modulator compounds, such as iminobiotinyl-$A\beta_{6-20}$ (PPI-267), cholyl-$A\beta_{16-20}$ (PPI-350) and cholyl-$A\beta_{16-20}$-amide (PPI-319).

Additionally, the β-amyloid modulator compounds have been demonstrated to reduce the neurotoxicity of preformed $A\beta_{1-40}$ aggregates. In these experiments, $A\beta_{1-40}$ aggregates were preformed by incubation of the monomers in the absence of any modulators. The modulator compound was then incubated with the preformed $A\beta_{1-40}$ aggregates for 24 hours at 37° C., after which time the β-AP/modulator solution was collected and its neurotoxicity evaluated as described above. Incubation of preformed $A\beta_{1-40}$ aggregates with the modulator compound prior to applying the solution to neuronal cells resulted in a decrease in the neurotoxicity of the $A\beta_{1-40}$ solution. These results suggest that the modulator can either bind to Aβ fibrils or soluble aggregate and modulate their inherent neurotoxicity or that the modulator can perturb the equilibrium between monomeric and aggregated forms of $A\beta_{1-40}$ in favor of the non-neurotoxic form.

EXAMPLE 11

Characterization of Additional β-Amyloid Modulator Compounds

In this example, additional modulator compounds designed based upon amino acids 17–20 of Aβ, LVFF; SEQ ID NO:12 (identified in Example 9), were prepared and analyzed to further delineate the structural features necessary for inhibition of β-amyloid aggregation. Types of compounds analyzed included ones having only three amino acid residues of an Aβ aggregation core domain, compounds in which the amino acid residues of an Aβ aggregation core domain were rearranged or in which amino acid substitutions had been made, compounds modified with a carboxy-terminal modifying group and compounds in which the modifying group had been derivatized. Abbreviations used in this example are: h- (free amino terminus), -oh (free

carboxylic acid terminus), -nh_2 (amide terminus), CA (cholyl, the acyl portion of cholic acid), NANA (N-acetyl neuraminyl), IB (iminobiotinyl), βA (β-alanyl), DA (D-alanyl), Adp (aminoethyldibenzofuranylpropanoic acid), Aic (3-(O-aminoethyl-iso)-cholyl, a derivative of cholic acid), IY (iodotyrosyl), o-methyl (carboxy-terminal methyl ester), N-me (N-methyl peptide bond), DeoxyCA (deoxycholyl) and LithoCA (lithocholyl).

Modulator compounds having an Aic modifying group at either the amino- or carboxy-terminus (e.g., PPI-408 and PPI-418) were synthesized using known methods (see e.g., Wess, G. et al. (1993) *Tetrahedron Letters,* 34:817–822; Wess, G. et al. (1992) *Tetrahedron Letters* 33:195–198). Briefly, 3-iso-O-(2-aminoethyl)-cholic acid (3β-(2-aminoethoxy)-7α,12α-dihydroxy-5β-cholanoic acid) was converted to the FMOC-protected derivative using FMOC-OSu (the hydroxysuccinimide ester of the FMOC group, which is commercially available) to obtain a reagent that was used to introduce the cholic acid derivative into the compound. For N-terminal introduction of the cholic acid moiety, the FMOC-protected reagent was coupled to the N-terminal amino acid of a solid-phase peptide in the standard manner, followed by standard FMOC-deprotection conditions and subsequent cleavage from the resin, followed by HPLC purification. For C-terminal introduction of the cholic acid moiety, the FMOC-protected reagent was attached to 2-chlorotrityl chloride resin in the standard manner. This amino acyl derivatized resin was then used in the standard manner to synthesize the complete modified peptide.

The modulators were evaluated in the nucleation and seeded extension assays described in Example 6 and the results are summarized below in Table V. The change in lag time (ΔLag) is presented as the ratio of the lag time observed in the presence of the test compound to the lag time of the control. Data are reported for assays in the presence of 100 mole % inhibitor relative to the concentration of $A\beta_{1-40}$, except for PPI-315, PPI-348, PPI-380, PPI-407 and PPI-418, for which the data is reported in the presence of 33 mole % inhibitor. Inhibition (% $I_{nucl'n}$) is listed as the percent reduction in the maximum observed turbidity in the control at the end of the assay time period. Inhibition in the extension assay (% $I_{ext'n}$) is listed as the percent reduction of thioflavin-T fluorescence of β-structure in the presence of 25 mole % inhibitor. Compounds with a % $I_{nucl'n}$ of at least 30% are highlighted in bold.

TABLE V

Ref. #	N-Term. Mod.	Peptide	C-Term. Mod.	ΔLag	% $I_{nucl'n}$	% $I_{ext'n}$
PPI-293	CA	----	-oh	1.0	0	ND*
PPI-315	CA	HQKLVFF	-nh_2	1.1	5**	ND
PPI-316	NANA	HQKLVFF	-nh_2	1.5	−15	ND
PPI-319	CA	KLVFF	-nh_2	5.4	70	52
PPI-339	CA	HDSGY	-nh_2	1.1	−18	ND
PPI-348	CA	HQKLVFF	-oh	2.0	70**	ND
PPI-349	CA	QKLVFF	-oh	>5	100	56
PPI-350	CA	KLVFF	-oh	1.8	72	11
PPI-365	CA	AAAAA	-oh	0.8	−7	0
PPI-366	CA	QKLVF	-oh	3.1	−23	ND
PPI-368	CA	LVFFA	-oh	>5	100	91
PPI-369	CA	KLTFF	oh	1.1	−16	44
PPI-370	CA	KLVAF	-oh	2.6	73	31
PPI-371	CA	KLVFF(βA)	-oh	2.5	76	80
PPI-372	CA	FKFVL	-oh	0.8	45	37
PPI-373	NANA	KLVFF	-oh	0.9	16	8
PPI-374	IB	KLVFF	-oh	1.3	86	0
PPI-375	CA	KTVFF	-oh	1.2	18	21

TABLE V-continued

Ref. #	N-Term. Mod.	Peptide	C-Term. Mod.	ΔLag	% $I_{nucl'n}$	% $I_{ext'n}$
PPI-377	h-	KLVFF	-oh	1.1	0	8
PPI-379	CA	LVFFAE	-oh	1.4	55	16
PPI-380	CA	LVFF	-oh	1.8	72**	51
PPI-381	CA	LVFF(DA)	-oh	2.3	56	11
PPI-382	CA	LVFFA	-nh_2	1.0	−200	91
PPI-383	h-DDIIL-(Adp)	VFF	-oh	0.4	14	0
PPI-386	h-	LVFFA	-oh	1.0	15	11
PPI-387	h-	KLVFF	-nh_2	1.3	−9	39
PPI-388	CA	AVFFA	-oh	1.4	68	44
PPI-389	CA	LAFFA	-oh	1.5	47	66
PPI-390	CA	LVAFA	-oh	2.7	25	0
PPI-392	CA	VFFA	-oh	2.0	76	10
PPI-393	CA	LVF	-oh	1.3	1	0
PPI-394	CA	VFF	-oh	1.8	55	0
PPI-395	CA	FFA	-oh	1.0	51	6
PPI-396	CA	LV(IY)FA	-oh	>5	100	71
PPI-401	CA	LVFFA	-o-methyl	ND	ND	0
PPI-405	h-	LVFFA	-nh_2	1.3	11	70
PPI-407	CA	LVFFK	-oh	>5	100**	85
PPI-408	h-	LVFFA	(Aic)-oh	3.5	46	3
PPI-418	h-(Aic)	LVFFA	-oh	>5	100**	87
PPI-426	CA	FFVLA	-oh	>5	100	89
PPI-391	CA	LVFAA	-oh	1.6	40	ND
PPI-397	CA	LVF(IY)A	-oh	>5	95	ND
PPI-400	CA	AVAFA	-oh	1.0	−15	ND
PPI-403	***	HQKLVFF	-oh	1.4	−75	0
PPI-404	****	LKLVFF	-oh	1.8	−29	7
PPI-424	DeoxyCA	LVFFA	-oh	3.0	−114	82
PPI-425	LithoCA	LVFFA	-oh	2.8	−229	0
PPI-428	CA	FF	-oh	1.7	−78	15
PPI-429	CA	FFV	-oh	2.2	−33	7
PPI-430	CA	FFVL	-oh	4.1	33	75
PPI-433	CA	LVFFA (all D amino acids)	-oh	2.8	27	ND
PPI-435	t-Boc	LVFFA	-oh	3.0	−5	ND
PPI-438	CA	GFF	-oh	1.0	0	ND

*ND = not done
**33 mol %
***h-DDIII(N-Me-Val)DLL(Adp)
****h-DDII(N-Me-Leu)VEH(Adp)

Certain compounds shown in Table V (PPI-319, PPI-349, PPI-350, PPI-368 and PPI-426) also were tested in neurotoxicity assays such as those described in Example 10. For each compound, the delay of the appearance of neurotoxicity relative to control coincided with the delay in the time at which polymerization of Aβ began in the nucleation assays. This correlation between the prevention of formation of neurotoxic Aβ species and the prevention of polymerization of Aβ was consistently observed for all compounds tested.

The results shown in Table V demonstrate that at an effective modulator compound can comprise as few as three Aβ amino acids residues (see PPI-394, comprising the amino acid sequence VFF, which corresponds to $A\beta_{18-20}$, and PPI-395, comprising the amino acid sequence FFA, which corresponds to $A\beta_{19-21}$). The results also demonstrate that a modulator compound having a modulating group at its carboxy-terminus is effective at inhibiting Aβ aggregation (see PPI-408, modified at its C-terminus with Aic). Still further, the results demonstrate that the cholyl group, as a modulating group, can be manipulated while maintaining the inhibitory activity of the compounds (see PPI-408 and PPI-418, both of which comprise the cholyl derivative Aic). The free amino group of the Aic derivative of cholic acid represents a position at which a chelation group for ^{99m}Tc can be introduced, e.g., to create a diagnostic agent. Additionally, the ability to substitute iodotyrosyl for phenylalanine at position 19 or 20 of the Aβ sequence (see PPI-396

and PPI-397) while maintaining the ability of the compound to inhibit Aβ aggregation indicates that the compound could be labeled with radioactive iodine, e.g., to create a diagnostic agent, without loss of the inhibitory activity of the compound.

Finally, compounds with inhibitory activity were created using Aβ derived amino acids but wherein the amino acid sequence was rearranged or had a substitution with a non-Aβ-derived amino acid. Examples of such compounds include PPI-426, in which the sequence of $A\beta_{17-21}$ (LVFFA SEQ ID NO:11) has been rearranged (FFVLA SEQ ID NO:21), PPI-372, in which the sequence of $A\beta_{16-20}$ (KLVFF SEQ ID NO:10) has been rearranged (FKFVL SEQ ID NO:29), and PPI-388, -389 and -390, in which the sequence of $A\beta_{17-21}$ (LVFFA SEQ ID NO:11) has been substituted at position 17, 18 or 19, respectively, with an alanine residue (AVFFA (SEQ ID NO:25) for PPI-388, LAFFA (SEQ ID NO:13) for PPI-389 and LVAFA (SEQ ID NO:33) for PPI-390). The inhibitory activity of these compounds indicate that the presence in the compound of an amino acid sequence directly corresponding to a portion of Aβ is not essential for inhibitory activity, but rather suggests that maintenance of the hydrophobic nature of this core region, by inclusion of amino acid residues such as phenylalanine, valine, leucine, regardless of their precise order, can be sufficient for inhibition of Aβ aggregation.

EXAMPLE 12

Characterization of β-Amyloid Modulator Compounds Comprising an Unmodified β-Amyloid Peptide

To examine the ability of unmodified Aβ peptides to modulate aggregation of natural β-AP, a series of Aβ peptides having amino- and/or carboxy terminal deletions as compared to $A\beta_{1-40}$, or having internal amino acids deleted (i.e., noncontiguous peptides), were prepared. One peptide (PPI-220) had additional, non-Aβ-derived amino acid residues at its amino-terminus. These peptides all had a free amino group at the amino-terminus and a free carboxylic acid at the carboxy-terminus. These unmodified peptides were evaluated in assays as described in Example 7. The results are summarized below in Table VI, wherein the data is presented as described in Example 7. Compounds exhibiting at least a 1.5 fold increase in lag time are highlighted in bold.

TABLE VI

Reference #	Aβ Peptide	Mole %	Fold Increase in Lag Time	% Inhibition of Plateau
PPI-226	$A\beta_{6-20}$	100	1.66	76
PPI-227	$A\beta_{11-25}$	100	~1	47
PPI-228	$A\beta_{16-30}$	100	>4.5	100
PPI-229	$A\beta_{21-35}$	100	~1	~0
PPI-230	$A\beta_{26-40}$	100	0.8	~0
PPI-231	$A\beta_{1-15}$	100	~1	18
PPI-247	$A\beta_{1-30,\ 36-40}$ (Δ31-35)	100	~1	~0
PPI-248	$A\beta_{1-25,\ 31-40}$ (Δ26-30)	100	1.58	~0
PPI-249	$A\beta_{1-20,\ 36-40}$ (Δ21-25)	100	2.37	~0
PPI-250	$A\beta_{1-15,\ 21-40}$ (Δ16-20)	100	1.55	~0
PPI-251	$A\beta_{1-10,\ 16-40}$ (Δ11-15)	100	~1.2	~0
PPI-252	$A\beta_{1-5,\ 11-40}$ (Δ6-10)	100	1.9	33
PPI-253	$A\beta_{6-40}$	100	1.9	~0
PPI-220	EEVVHHHHQQ-$A\beta_{16-40}$	100	>4	100

The results shown in Table VI demonstrate that limited portions of the Aβ sequence can have a significant inhibitory effect on natural β-AP aggregation even when the peptide is not modified by a modifying group. Preferred unmodified peptides are $A\beta_{6-20}$ (PPI-226), $A\beta_{16-30}$ (PPI-228), $A\beta_{1-20,\ 26-40}$ (PPI-249) and EEVVHHHHQQ-$A\beta_{16-20}$ (PPI-220), the amino acid sequences of which are shown in SEQ ID NOs:4, 14, 15, and 16, respectively.

Forming part of this disclosure is the appended Sequence Listing, the contents of which are summarize in the Table below.

SEQ ID NO:	Amino Acids	Peptide Sequence
1	43 amino acids	$A\beta_{1-43}$
2	103 amino acids	APP C-terminus
3	43 amino acids	$A\beta_{1-43}$ (19, 20 mutated)
4	HDSGYEVHHQKLVFF	$A\beta_{6-20}$
5	HQKLVFFA	$A\beta_{14-21}$
6	HQKLVFF	$A\beta_{14-20}$
7	QKLVFFA	$A\beta_{15-21}$
8	QKLVFF	$A\beta_{15-20}$
9	KLVFFA	$A\beta_{16-21}$
10	KLVFF	Aβ16-20
11	LVFFA	$A\beta_{17-21}$
12	LVFF	$A\beta_{17-20}$
13	LAFFA	$A\beta_{17-21}$ (V_{18}→A)
14	KLVFFAEDVGSNKGA	$A\beta_{16-30}$
15	35 amino acids	$A\beta_{1-20,\ 26-40}$
16	35 amino acids	EEVVHHHHQQ-βAP_{16-40}
17	AGAAAAGA	PrP peptide
18	AILSS	amylin peptide
19	VFF	$A\beta_{18-20}$
20	FFA	$A\beta_{19-21}$
21	FFVLA	$A\beta_{17-21}$ (scrambled)
22	LVFFK	$A\beta_{17-21}$ (A_{21}→K)
23	LV(IY)FA	$A\beta_{17-21}$ (F_{19}→IY)
24	VFFA	$A\beta_{18-21}$
25	AVFFA	$A\beta_{17-21}$ (L_{17}→A)
26	LVF(IY)A	$A\beta_{17-21}$ (F_{20}→IY)
27	LVFFAE	$A\beta_{17-22}$
28	FFVL	$A\beta_{17-20}$ (scrambled)
29	FKFVL	$A\beta_{16-20}$ (scrambled)
30	KLVAF	$A\beta_{16-20}$ (F_{19}→A)
31	KLVFF(βA)	$A\beta_{16-21}$ (A_{21}→βA)
32	LVFF(DA)	$A\beta_{17-21}$ (A_{21}→DA)

EQUIVALENTS

Those skilled in the art will recognize, or be able to ascertain using no more than routine experimentation, many equivalents to the specific embodiments of the invention described herein. Such equivalents are intended to be encompassed by the following claims.

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i i i) NUMBER OF SEQUENCES: 40

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 43 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: peptide

(v) FRAGMENT TYPE: internal

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:1:

```
Asp Ala Glu Phe Arg His Asp Ser Gly Tyr Glu Val His His Gln Lys
 1               5                  10                  15

Leu Val Phe Phe Ala Glu Asp Val Gly Ser Asn Lys Gly Ala Ile Ile
            20                  25                  30

Gly Leu Met Val Gly Gly Val Val Ile Ala Thr
        35                  40
```

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 103 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: peptide

(v) FRAGMENT TYPE: internal

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:2:

```
Glu Val Lys Met Asp Ala Glu Phe Arg His Asp Ser Gly Tyr Glu Val
 1               5                  10                  15

His His Gln Lys Leu Val Phe Phe Ala Glu Asp Val Gly Ser Asn Lys
            20                  25                  30

Gly Ala Ile Ile Gly Leu Met Val Gly Gly Val Val Ile Ala Thr Val
        35                  40                  45

Ile Val Ile Thr Leu Val Met Leu Lys Lys Lys Gln Tyr Thr Ser Ile
    50                  55                  60

His His Gly Val Val Glu Val Asp Ala Ala Val Thr Pro Glu Glu Arg
65                  70                  75                  80

His Leu Ser Lys Met Gln Gln Asn Gly Tyr Glu Asn Pro Thr Tyr Lys
                85                  90                  95

Phe Phe Glu Gln Met Gln Asn
            100
```

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 43 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: peptide

(v) FRAGMENT TYPE: internal

(i x) FEATURE:
(A) NAME/KEY: Modified site
(B) LOCATION: 19

-continued

(D) OTHER INFORMATION: /note= Xaa is a hydrophobic amino acid

(ix) FEATURE:
(A) NAME/KEY: Modified site
(B) LOCATION: 20
(D) OTHER INFORMATION: /note= Xaa is a hydrophobic amino acid

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

Asp Ala Glu Phe Arg His Asp Ser Gly Tyr Glu Val His His Gln Lys
1 5 10 15

Leu Val Xaa Xaa Ala Glu Asp Val Gly Ser Asn Lys Gly Ala Ile Ile
20 25 30

Gly Leu Met Val Gly Gly Val Val Ile Ala Thr
35 40

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 15 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

His Asp Ser Gly Tyr Glu Val His His Gln Lys Leu Val Phe Phe
5 10 15

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 8 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

His Gln Lys Leu Val Phe Phe Ala
5

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 7 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

His Gln Lys Leu Val Phe Phe
5

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 7 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

Gln Lys Leu Val Phe Phe Ala
5

-continued

(2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 6 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: peptide

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:8:

Gln Lys Leu Val Phe Phe
5

(2) INFORMATION FOR SEQ ID NO:9:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 6 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: peptide

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:9:

Lys Leu Val Phe Phe Ala
5

(2) INFORMATION FOR SEQ ID NO:10:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 5 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: peptide

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:10:

Lys Leu Val Phe Phe
5

(2) INFORMATION FOR SEQ ID NO:11:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 5 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: peptide

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:11:

Leu Val Phe Phe Ala
5

(2) INFORMATION FOR SEQ ID NO:12:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 4 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: peptide

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:12:

Leu Val Phe Phe

(2) INFORMATION FOR SEQ ID NO:13:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 5 amino acids
(B) TYPE: amino acid

-continued

(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: peptide

(v) FRAGMENT TYPE: internal

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:13:

Leu Ala Phe Phe Ala
1 5

(2) INFORMATION FOR SEQ ID NO:14:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 15 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: peptide

(v) FRAGMENT TYPE: internal

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:14:

Lys Leu Val Phe Phe Ala Glu Asp Val Gly Ser Asn Lys Gly Ala
1 5 10 15

(2) INFORMATION FOR SEQ ID NO:15:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 35 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: peptide

(v) FRAGMENT TYPE: internal

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:15:

Asp Ala Glu Phe Arg His Asp Ser Gly Tyr Glu Val His His Gln Lys
1 5 10 15

Leu Val Phe Phe Ser Asn Lys Gly Ala Ile Ile Gly Leu Met Val Gly
20 25 30

Gly Val Val
35

(2) INFORMATION FOR SEQ ID NO:16:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 35 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: peptide

(v) FRAGMENT TYPE: internal

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:16:

Glu Glu Val Val His His His His Glu Gln Lys Leu Val Phe Phe Ala
1 5 10 15

Glu Asp Val Gly Ser Asn Lys Gly Ala Ile Ile Gly Leu Met Val Gly
20 25 30

Gly Val Val
35

(2) INFORMATION FOR SEQ ID NO:17:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 8 amino acids
(B) TYPE: amino acid

-continued

(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: peptide

(v) FRAGMENT TYPE: internal

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:17:

Ala Gly Ala Ala Ala Ala Gly Ala
1 5

(2) INFORMATION FOR SEQ ID NO:18:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 5 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: peptide

(v) FRAGMENT TYPE: internal

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:18:

Ala Ile Leu Ser Ser
1 5

(2) INFORMATION FOR SEQ ID NO:19:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 3 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: peptide

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:19:

Val Phe Phe
1

(2) INFORMATION FOR SEQ ID NO:20:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 3 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: peptide

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:20:

Phe Phe Ala
1

(2) INFORMATION FOR SEQ ID NO:21:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 5 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: peptide

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:21:

Phe Phe Val Leu Ala
1 5

(2) INFORMATION FOR SEQ ID NO:22:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 5 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

-continued

(i i) MOLECULE TYPE: peptide

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:22:

Leu Val Phe Phe Lys
1 5

(2) INFORMATION FOR SEQ ID NO:23:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 5 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: peptide

(i x) FEATURE:
(A) NAME/KEY: Modified site
(B) LOCATION: 3
(D) OTHER INFORMATION: /note= Xaa is iodotyrosyl

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:23:

Leu Val Xaa Phe Ala
1 5

(2) INFORMATION FOR SEQ ID NO:24:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 4 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: peptide

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:24:

Val Phe Phe Ala
1

(2) INFORMATION FOR SEQ ID NO:25:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 5 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: peptide

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:25:

Ala Val Phe Phe Ala
1 5

(2) INFORMATION FOR SEQ ID NO:26:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 5 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: peptide

(i x) FEATURE:
(A) NAME/KEY: Modified site
(B) LOCATION: 4
(D) OTHER INFORMATION: /note= Xaa is iodotyrosyl

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:26:

Leu Val Phe Xaa Ala
1

(2) INFORMATION FOR SEQ ID NO:27:

-continued

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 6 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: peptide

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:27:

Leu Val Phe Phe Ala Glu
1 5

(2) INFORMATION FOR SEQ ID NO:28:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 4 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: peptide

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:28:

Phe Phe Val Leu
1 5

(2) INFORMATION FOR SEQ ID NO:29:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 5 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: peptide

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:29:

Phe Lys Phe Val Leu
1 5

(2) INFORMATION FOR SEQ ID NO:30:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 5 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: peptide

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:30:

Lys Leu Val Ala Phe
1 5

(2) INFORMATION FOR SEQ ID NO:31:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 6 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: peptide

(i x) FEATURE:
(A) NAME/KEY: Modified site
(B) LOCATION: 6
(D) OTHER INFORMATION: /note= Xaa is beta-alanyl

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:31:

Lys Leu Val Phe Phe Xaa
1

(2) INFORMATION FOR SEQ ID NO:32:

-continued

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 5 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: peptide

(i x) FEATURE:
(A) NAME/KEY: Modified site
(B) LOCATION: 5
(D) OTHER INFORMATION: /note= Xaa is D-alanyl

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:32:

Leu Val Phe Phe Xaa
1

(2) INFORMATION FOR SEQ ID NO:33:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 5 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: peptide

(v) FRAGMENT TYPE: internal

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:33:

Leu Val Ala Phe Ala
1 5

(2) INFORMATION FOR SEQ ID NO:34:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 5 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: peptide

(v) FRAGMENT TYPE: internal

(i x) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 5
(D) OTHER INFORMATION: /note=aminoethyldibenzofuranyl-
proprionic acid modification

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:34:

Asp Asp Ile Ile Leu
1 5

(2) INFORMATION FOR SEQ ID NO:35:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 5 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: peptide

(v) FRAGMENT TYPE: internal

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:35:

Ala Ala Ala Ala Ala
1 5

(2) INFORMATION FOR SEQ ID NO:36:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 35 amino acids

-continued

(B) TYPE: amino acid
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: peptide

(v) FRAGMENT TYPE: internal

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:36:

```
His Asp Ser Gly Tyr Glu Val His His Gln Lys Leu Val Phe Phe Ala
 1               5                  10                  15

Glu Asp Val Gly Ser Asn Lys Gly Ala Ile Ile Gly Leu Met Val Gly
            20                  25                  30

Gly Val Val
        35
```

(2) INFORMATION FOR SEQ ID NO:37:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 15 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: peptide

(v) FRAGMENT TYPE: internal

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:37:

```
Glu Val His His Gln Lys Leu Val Phe Phe Ala Glu Asp Val Gly
 1               5                  10                  15
```

(2) INFORMATION FOR SEQ ID NO:38:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 35 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: peptide

(v) FRAGMENT TYPE: internal

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:38:

```
Asp Ala Glu Phe Arg His Asp Ser Gly Tyr Glu Val His His Gln Lys
 1               5                  10                  15

Leu Val Phe Phe Ala Glu Asp Val Gly Ile Ile Gly Leu Met Val Gly
            20                  25                  30

Gly Val Val
        35
```

(2) INFORMATION FOR SEQ ID NO:39:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 35 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: peptide

(v) FRAGMENT TYPE: internal

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:39:

```
Asp Ala Glu Phe Arg His Asp Ser Gly Tyr Glu Val His His Gln Ala
 1               5                  10                  15

Glu Asp Val Gly Ser Asn Lys Gly Ala Ile Ile Gly Leu Met Val Gly
            20                  25                  30

Gly Val Val
        35
```

-continued

(2) INFORMATION FOR SEQ ID NO:40:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 35 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(i i) MOLECULE TYPE: peptide

(v) FRAGMENT TYPE: internal

(x i) SEQUENCE DESCRIPTION: SEQ ID NO:40:

Asp Ala Glu Phe Arg Glu Val His His Gln Lys Leu Val Phe Phe Ala
1 5 10 15

Glu Asp Val Gly Ser Asn Lys Gly Ala Ile Ile Gly Leu Met Val Gly
20 25 30

Gly Val Val
35

We claim:

1. A β-amyloid peptide compound having a structure: βAP_{6-20} (SEQ ID NO:4).

2. A β-amyloid peptide compound having a structure: βAP_{16-30} (SEQ ID NO:14).

3. A β-amyloid peptide compound having a structure: $\beta AP_{1-20,\ 26-40}$ (SEQ ID NO:15).

4. A β-amyloid peptide compound having a structure: EEVVHHHHQQ-βAP_{16-40} (SEQ ID NO:16).

5. A β-amyloid peptide compound having a structure: $A\beta_{6-40}$ (SEQ ID NO:36).

6. A β-amyloid peptide compound having a structure: $A\beta_{11-25}$ (SEQ ID NO:37).

7. A β-amyloid peptide compound having a structure: $A\beta_{1-25,\ 31-40}$ (Δ26–30) (SEQ ID NO:38).

8. A β-amyloid peptide compound having a structure: $A\beta_{1-15,\ 21-40}$ (Δ16–20) (SEQ ID NO:39).

9. A β-amyloid peptide compound having a structure: $A\beta_{1-5,\ 11-40}$ (Δ6–10) (SEQ ID NO:40).

10. A composition comprising the compound of any one of claims 1–9 and a pharmaceutically acceptable carrier.

* * * * *

INTRODUCTION

Protein-protein interactions are intrinsic to virtually every cellular process. Any listing of major research topics in biology—for example, DNA replication, transcription, translation, splicing, secretion, cell cycle control, signal transduction, and intermediary metabolism—is also a listing of processes in which protein complexes have been implicated as essential components. In consequence, the analysis of the proteins in these complexes is no longer the exclusive domain of biochemists; geneticists, cell biologists, developmental biologists, molecular biologists, and biophysicists have by necessity all gotten into the act. We attempt in this review to summarize both classical and recent methods to identify proteins that interact and to assess the strengths of these interactions.

Proteins that are composed of more than one subunit are found in many different classes of proteins. Some of the best-characterized multisubunit proteins are those that, as originally purified, contained two or more different components. These include classical proteins such as hemoglobin, tryptophan synthetase, aspartate transcarbamylase, core RNA polymerase, Qβ-replicase, and glycyl-tRNA synthetase. Since these proteins purified as multisubunit complexes, their protein-protein interactions were self-evident.

Other well-known examples of multisubunit proteins include much more complicated assemblies of polypeptides. These include metabolic enzymes such as the pyruvate dehydrogenase and α-ketoglutarate dehydrogenase complexes, the DNA replication complex of *Escherichia coli* and other organisms, the bacterial flagellar apparatus, the nuclear pore complex, and the tail assembly of bacteriophage T4. Also included in this group are ribonucleoprotein complexes, such as the signal recognition particle of the glycosylation pathway, small nuclear ribonucleoproteins of the spliceosome, and the ribosome itself. Although some of the subunits of these protein complexes are not tightly bound, activity is associated with a large structure that in many cases is called a protein machine (5).

There are also a large number of transient protein-protein interactions, which in turn control a large number of cellular processes. All modifications of proteins necessarily involve such transient protein-protein interactions. These include the interactions of protein kinases, protein phosphatases, glycosyl transferases, acyl transferases, proteases, etc., with their substrate proteins. Such protein-modifying enzymes encompass a large number of protein-protein interactions in the cell and regulate all manner of fundamental processes such as cell growth, cell cycle, metabolic pathways, and signal transduction. Transient protein-protein interactions are also involved in the recruitment and assembly of the transcription complex to specific promoters, the transport of proteins across membranes, the folding of native proteins catalyzed by chaperonins, individual steps of the translation cycle, and the breakdown and re-formation of subcellular structures during the cell cycle (such as the cytoplasmic microtubules, the spindle apparatus, nuclear lamina, and the nuclear pore complex). Transient complexes are much more difficult to study, because the proteins or conditions responsible for the transient reaction have to be identified first. Part of the goal of this review is to describe recent methods and developments that have allowed their identification and characterization.

Protein-protein interactions can have a number of different measurable effects. First, they can alter the kinetic properties of proteins. This can be reflected in altered binding of substrates, altered catalysis, or (as first enunciated by Monod et al. [153]) altered allosteric properties of the complex. Thus, the interaction of proliferating-cell nuclear antigen with DNA polymerase δ alters the processivity of the polymerase (174), the interaction of succinate thiokinase and α-ketoglutarate dehydrogenase lowers the K_m for succinyl coenzyme A by 30-fold (171), and the cooperative binding of oxygen to hemoglobin and the allosteric regulation of aspartate transcarbamylase are regulated by interactions of the protomers. Second, protein-protein interactions are one common mechanism to allow for substrate channeling. The paradigm for this type of complex is tryptophan synthetase from *Neurospora crassa*. It is a complex of two subunits, each of which carries out one of the two steps of reaction (formation of indole from indole 3-glycerol phosphate, followed by conversion of indole to tryptophan). The intermediate indole is noncovalently bound, but it is preferentially channeled to form tryptophan (241). Many similar examples of metabolic channeling have been demonstrated, both between different subunits of a complex and between different domains of a single multifunctional polypeptide (see reference 208 for a review). Third, protein-protein interactions can result in the formation of a new binding site. Thus, an ADP site forms at the interface of the α and β subunits of *Escherichia coli* F_1-ATPase (228), yeast hexokinase binds one ATP molecule at the interface of the asymmetric homodimer (209), and phosphofructokinase from *Bacillus stearothermophilus* binds both fructose 6-phosphate and ADP at the interface between subunits (60). Fourth, protein-protein interactions can inactivate a protein; this is the case with the interaction of phage P22 repressor with its antirepressor (213), with the interaction of trypsin with trypsin inhibitor (221), and with the interaction of

phage T7 gene 1.2 protein with *E. coli* dGTP triphosphohydrolase (156). Fifth, protein-protein interactions can change the specificity of a protein for its substrate; thus, the interaction of lactalbumin with lactose synthase lowers the K_m for glucose by 1,000-fold (95), and the interaction of transcription factors with RNA polymerase directs the polymerase to different promoters.

Klotz et al. (116) enumerated four advantages of multisubunit proteins relative to a single large protein with multiple sites. First, it is much more economical to build proteins from simpler subunits than to require multiple copies of the coding information to synthesize oligomers. Thus, for example, actin filaments and virus coats are much more simply assembled from monomers than by translation of a large polyprotein of repeated domains. Similarly, it is much more convenient to have one gene encoding a protein with different interacting partners, such as some of the eukaryotic RNA polymerase subunits, than to have the gene for that subunit reiterated for each different polymerase. Second, translation of large proteins can cause a significant increase in errors in translation; if such errors cause a lack of activity, they are much more economically eliminated by preventing assembly of that subunit into the complex than by eliminating the whole protein. Third, multisubunit assemblies allow for synthesis at one locale, followed by diffusion and assembly at another locale; this allows for both faster diffusion (since the monomers are smaller) and compartmentalization of activity (if assembly is required for activity). Fourth, homooligomeric proteins, if they have an advantage over monomers, are easily selected in evolution if the oligomers interact in an antiparallel arrangement; in this case, a single-amino-acid change that increases interaction potential has effects at two such sites.

Another advantage of multisubunit complexes is the ability to use different combinations of subunits to alter the magnitude or type of response. Thus, for example, adult hemoglobin ($\alpha2\beta2$) and fetal hemoglobin ($\alpha2\gamma2$) are each composed of heterooligomers with a common α subunit; differences in the binding of oxygen in these hemoglobins allow oxygen to be readily passed from mother to fetus. Other examples include the oligomerization of Jun with Fos or with itself, which results in distinct activities in transcription because the different dimers bend DNA in opposite directions (114); the interaction of TATA-binding protein with the transcription apparatus of RNA polymerase I, II, or III, in which TATA-binding protein plays different roles (235); the interactions of microtubules with the large set of proteins to which they bind (113), not all of which bind at the same time; the interaction of different transcription factors with core RNA polymerase in both eukaryotes and prokaryotes to direct transcription of different genes; and the interaction of retinoblastoma (Rb) protein with viral oncoproteins and other cellular proteins (31, 32).

Protein-protein interactions may be mediated at one extreme by a small region of one protein fitting into a cleft in another protein and at another extreme by two surfaces interacting over a large area. Examples of the first case include the large class of protein-protein interactions that involve a domain of a protein interacting tightly with a small peptide. The paradigm for this type of interaction is that of specific Src homology 2 (SH2) domains with specific small peptides containing a phosphotyrosyl residue. This interaction occurs with a dissociation constant as low as nM and is due to a specific binding pocket in SH2 domains not unlike a classical substrate-binding pocket (64, 205, 224, 225). Many other examples of domains that bind small peptides with affinities in the nanomolar to molar range have been described. The paradigm for the second case, i.e., surfaces that interact with each other over large areas, is that of the leucine zipper, in which a stretch of α-helix forms a surface that fits almost perfectly with another α-helix from another subunit protein (59, 161; also see reference 4). Binding also occurs in the nanomolar range for such interactions (196). Other interactions may occur through intermediate-sized complementary surfaces.

It is evident that protein-protein interactions are much more widespread than once suspected, and the degree of regulation that they confer is large. To properly understand their significance in the cell, one needs to identify the different interactions, understand the extent to which they take place in the cell, and determine the consequences of the interaction. This review is intended to supply an overview of three aspects of protein-protein interactions. First, we briefly describe a number of physical, molecular biological, and genetic approaches that have been used to detect protein-protein interactions. Second, we describe several experimental approaches that have been used to evaluate the strength of protein-protein interactions. Third, we describe three well-characterized domains that are responsible for protein-protein interactions in a number of different proteins. As the literature on this topic is vast, we have not attempted to conduct an exhaustive review. Rather, we hope that this article serves as a journeyman's guide to protein-protein interactions.

The first and still the most comprehensive review on protein-protein interactions is that of Klotz et al. (116). This review contains a survey of the subunit composition and binding energies of all oligomeric proteins that had been identified at the time, as well as a discussion of the geometry of interactions and an excellent discussion of the influence of binding constants, concentrations, and cooperativity parameters on the population of oligomers. A good discussion of channeling and compartmentation is found in the monograph by Friedrich on quaternary structure (70) and the article by Srere (208). The review by Eisenstein and Schachman (57) contains an interesting discussion of the functional roles of subunits of oligomeric proteins and of approaches used to determine whether the monomers of oligomeric proteins are active. Also of interest is the discussion of proteins as machines (5) and a discussion of protein size and composition (78).

PHYSICAL METHODS TO SELECT AND DETECT PROTEINS THAT BIND ANOTHER PROTEIN

Protein Affinity Chromatography

A protein can be covalently coupled to a matrix such as Sepharose under controlled conditions and used to select ligand proteins that bind and are retained from an appropriate extract. Most proteins pass through such columns or are readily washed off under low-salt conditions; proteins that are retained can then be eluted by high-salt solutions, cofactors, chaotropic solvents, or sodium dodecyl sulfate (SDS) (Fig. 1). If the extract is labeled in vivo before the experiment, there are two distinct advantages: labeled proteins can be detected with high sensitivity, and unlabeled polypeptides derived from the covalently bound protein can be ignored (these might be either proteolytic fragments of the covalently bound protein or subunits of the protein which are not themselves covalently bound). This method was first used 20 years ago to detect phage and host proteins that interacted with different forms of *E. coli* RNA polymerase (177). Proteins that were retained by an RNA polymerase-agarose column (which was shown to be enzymatically active) but not by a control column coupled with bovine serum albumin were judged as interacting candidates. The interactions were substantiated in two ways. First, the

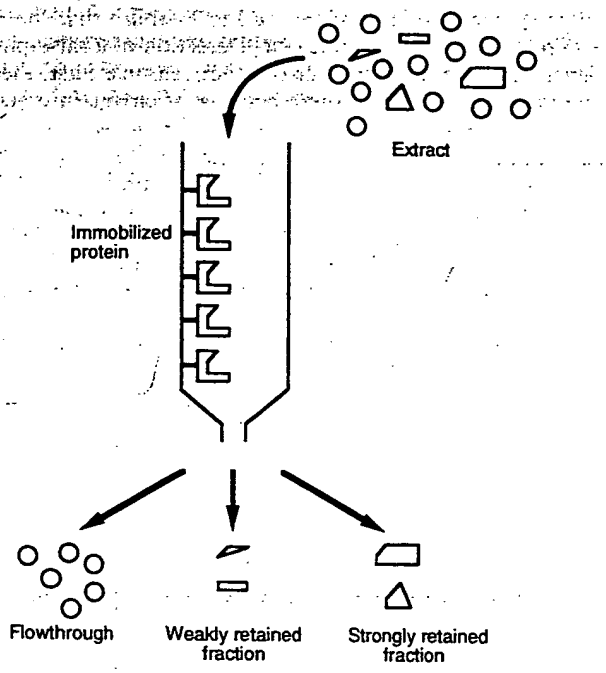


FIG. 1. Protein affinity chromatography. Extract proteins are passed over a column containing immobilized protein. Proteins that do not bind flow through the column, and ligand proteins that bind are retained. Strongly retained proteins have more contacts with the immobilized protein than do those that are weakly retained.

interaction of T7 0.3 protein with RNA polymerase was confirmed by coimmunoprecipitation of the 0.3 protein with RNA polymerase antibody. Second, the interaction of T4 proteins with RNA polymerase was shown to depend on the form of RNA polymerase on the column: one T4 protein interacted with core RNA polymerase and T4-modified RNA polymerase but not with RNA polymerase holoenzyme, and another interacted only with the T4-modified polymerase. The phage proteins that bound RNA polymerase were identified by their absence in appropriate T4 and T7 mutants.

Similar methods have been used, particularly by the laboratories of J. Greenblatt and B. Alberts, to identify many other protein-protein interactions. Two excellent reviews on the topic, which cover many of the details of coupling and a number of strategic considerations, have been published (69, 145).

Candidate proteins can be coupled directly to commercially available preactivated resins as described by Formosa et al. (69). Alternatively, they can be tethered noncovalently through high-affinity binding interactions. Thus, Beeckmans and Kanarek (14) demonstrated an interaction between fumarase and malate dehydrogenase by immobilizing the test enzyme with antibody bound to protein A-Sepharose, as well as by direct covalent coupling of the test enzyme to Sepharose. Some of the important considerations of a successful binding experiment are elaborated below.

Purity of the coupled protein and use of protein fusions. An essential requirement for a successful protein affinity chromatography experiment is pure protein; otherwise, any interacting protein that is detected might be binding to a contaminant in the preparation. Greenblatt and Li (80) did two experiments to establish that core RNA polymerase bound to NusA on the column rather than to a contaminant in the NusA preparation. First, they demonstrated that a fully active NusA variant protein, which presumably contained different amounts of various contaminants (since it eluted at different positions in columns used to purify it), still bound core RNA polymerase; second, they demonstrated by independent experiments that the complex contained equimolar amounts of NusA protein and core RNA polymerase.

The easiest way to obtain pure protein, if the gene is available, is through the use of protein fusions. Several such systems have been described; in each case, the protein of interest (or a domain of the protein) is fused to a protein or a domain that can be rapidly purified on the appropriate affinity resin. The most common such fusion contains glutathione *S*-transferase (GST), which can be purified on glutathione-agarose columns (202). Other fusions in common use include *Staphylococcus* protein A, which can be purified on columns bearing immunoglobulin G; oligohistidine-containing peptides, which can be purified on columns bearing Ni^{2+}; the maltose-binding protein, which can be purified on resins containing amylose; and dihydrofolate reductase, which can be purified on methotrexate columns. (Other common protein fusions which add an epitope for the influenza virus hemagglutinin [12CA5] or c-Myc are also in common use and are used most often for coimmunoprecipitation [see the section on immunoprecipitation, below].)

Purified fusion proteins are used in two ways to detect interactions on affinity columns. First, the protein is covalently coupled to the resins in the usual way, as was done by Mayer et al. (139) to detect a tyrosine-phosphorylated protein that bound to the SH2 domain of Abl tyrosine kinase and by Weng et al. (232) to demonstrate that the SH3 domain of c-Src binds paxillin. Second, the purified fusion proteins can be noncovalently bound to the beads and then mixed with an appropriate extract or protein. This was done by Zhang et al. (248) to demonstrate an interaction of the N-terminal portion of c-Raf with Ras, by Flynn et al. (68) to detect the binding of an actin filament-associated protein to Src-SH3/SH2, and by Hu et al. (99) to demonstrate the binding of the SH2 domain of the p85 subunit of phosphatidylinositol 3-kinase to two different growth factor receptors.

Influence of modification state. The interactions of many proteins with their target proteins often depends on the modification state of one or both of the proteins (mostly by phosphorylation). Thus, the recognition of Rb protein by the transcription factor E2F and by the transforming proteins simian virus 40 large T antigen, human papillomavirus-16 E7, and adenovirus E1A is more efficient with underphosphorylated than phosphorylated Rb (132, 133, 240). Conversely, SH2 domains of proteins, for example, recognize tyrosine phosphorylated substrates several orders of magnitude more efficiently than they do their nonphosphorylated counterparts (64). Protein-protein interactions that require a posttranslationally modified protein for interaction are not detected if the protein is purified by the use of expression vectors in cells in which the protein is not properly modified. A means to circumvent this problem is to use GST fusion vectors to express proteins in host cells more related to their origin. Thus, the interaction of bovine papillomavirus E5 oncoprotein with an α-adaptin-like molecule was confirmed by addition of beads to extracts of NIH 3T3 cells that were expressing the GST-E5 fusion (38). Similarly, a yeast GST vector that allows regulated expression of yeast GST fusion proteins has been described (148).

Retention f native structure of the coupled protein. Failure to detect an interacting protein can result from inactivation of the protein during coupling. Ideally, coupling would immobilize a protein or a complex by randomly tethering it to the matrix through one covalent bond. For example, binding of *E. coli* proteins to immobilized λ N protein occurred only when

the cyanogen bromide (CNBr)-activated residues on the matrix were partially inactivated before coupling; this was attributed to the large number of lysine residues in λ N protein and the generation of multiple (and denaturing) covalent bonds between λ N protein and the matrix if the concentration of CNBr-activated matrix sites was too high (80). Therefore, determining that the coupled protein has retained its native structure is an important control, when possible. With some proteins, such as RNA polymerase from *E. coli*, activity could be detected when the coupled protein was assayed on the matrix (177). With others, such as filamentous actin (F-actin) columns, the desired polymerized form was stabilized with phalloidin (or by chemical cross-linking), and the proteins that bound F-actin were shown not to bind monomeric actin (14). Similarly, microtubule columns were stabilized with taxol (113).

Native protein structure also depends on all subunits of a complex being present in the coupled resin. This can be assessed by SDS elution of a sample of the resin and comparison of the subunit composition of the eluted material with that of the starting material. In the case of *E. coli* RNA polymerase, all the components of the enzyme were still present (177). In the case of mammalian RNA polymerase II, one of the subunits did not reproducibly remain after coupling (206).

Concentration of the coupled protein. To detect interactions efficiently, the concentration of protein covalently bound to the column has to be well above the K_d of the interaction. Thus, for the detection of weak protein-protein interactions, the concentration of bound protein should be as high as possible. Weak interactions can be completely missed on columns with lower concentrations of coupled protein, even if they contain correspondingly larger amounts of resin to maintain the same total amount of bound protein (see the sections on importance of characterization of the binding interaction and on binding to immobilized proteins, below, for a discussion of this point).

Amount of extract applied. The amount of extract applied to the column can be critical for two opposing reasons. If too little extract is applied and the protein that binds is present at low concentration, too little protein will be retained to be detected, even if it binds with high affinity and is labeled with ^{35}S (see, for example, reference 206). Conversely, if too much protein is applied, competition among potential ligands may result in failure to detect minor species. This was observed by Miller and Alberts (144) in looking for minor protein species that interact with F-actin.

Other considerations. There are four distinct advantages of protein affinity chromatography as a technique for detecting protein-protein interactions. First, and most important, protein affinity chromatography is incredibly sensitive. With appropriate use (high concentrations of immobilized test protein), it can detect interactions with a binding constant as weak as 10^{-5} M (69) (see the section on binding to immobilized protein, below). This limit is within range of the weakest interaction likely to be physiologically relevant, which we estimate to be in the range of 10^{-3} M (see the section on limits of binding-constant considerations, below). Second, this technique tests all proteins in an extract equally; thus, extract proteins that are detected have successfully competed for the test protein with the rest of the population of proteins. Third, it is easy to examine both the domains of a protein and the critical residues within it that are responsible for a specific interaction, by preparing mutant derivatives (38, 216). Fourth, interactions that depend on a multisubunit tethered protein can be detected, unlike the case with protein blotting.

One potential problem derives from the very sensitivity of the technique. Since it detects interactions that are so weak, independent criteria must be used to establish that the interaction is physiologically relevant. Detection of a false-positive signal can arise for a number of other reasons. First, the protein may bind the test protein because of charge interactions; for this reason, it is desirable to use a control column with approximately the same ionic charges. Second, the proteins may interact through a second protein that interacts with the test protein; although interesting in itself, the interaction may not be direct. Third, the proteins may interact with high specificity even though they never encounter one another in the cell. The most famous example of this type is the high affinity of actin for DNase I (125).

For all of these reasons, the prudent course is to independently demonstrate the interaction in vitro or, if possible, in vivo. Cosedimentation was used to confirm the interaction of RAP 72 (now known as RAP 74) and RAP 30 with RNA polymerase II (206), NusA protein with core RNA polymerase (80), and NusB protein with ribosomal protein S10 (138). In other cases, more biological criteria were used. For example, antibodies were generated against many of the proteins that interacted with F-actin (but not monomeric G-actin) on columns, and these were used to demonstrate that more than 90% of the corresponding proteins were localized with an actin-like distribution during mitosis of *Drosophila* embryos at the syncytial blastoderm stage of development (144). The identification of three yeast actin-binding proteins was confirmed in three separate ways: one of the proteins was shown to correspond to the yeast analog of myosin by virtue of a shared epitope; another protein colocalized with actin cables and cortical actin patches, and overproduction of the third protein caused a reorganization of the actin cytoskeleton (53). In the identification of microtubule-associated proteins, two criteria were used to demonstrate the authenticity of the results (113). First, antibodies for 20 of the 24 candidate microtubule-associated proteins stain various parts of microtubule structures of *Drosophila* embryos during the cell cycle. Second, many (but not all) of the microtubule-associated proteins isolated on microtubule affinity columns are the same as those isolated by traditional cosedimentation methods of Vallee and Collins (219).

Failure to detect an interaction can occur for a number of technical reasons, described above. A false-negative result can arise for two additional reasons: the interacting protein may not be able to exchange with another protein to which it is binding, or the two proteins may not be able to interact both with each other and with the resin.

Protein affinity chromatography does not always yield answers corresponding to other approaches. For reasons that are unclear, a large number of proteins were detected by probing SDS-polyacrylamide gel electrophoresis (PAGE) gels with a GST fusion of the SH2 domain of Abl tyrosine kinase, but only a couple of proteins were detected on columns coupled with this protein (139). Similarly, a specific protein was detected on F-actin columns stabilized by suberimidate cross-linking but not with phalloidin (144). Finally, G-actin interacting proteins are very difficult to detect with columns of G-actin, although such columns bind DNase I; by contrast, DNase I columns can be used to detect such G-actin interactions (24).

Affinity Blotting

In a procedure analogous to th use of affinity columns, proteins can be fractionated by PAGE transferred to a nitrocellulose membrane, and identified by their ability to bind a protein, peptide, or other ligand. This method is similar to immunoblotting (Western blotting), which uses an antibody as

the probe. Complex mixtures of proteins, such as total-cell lysates, can be analyzed without any purification. Therefore, this method has been particularly useful for membrane proteins, such as cell surface receptors (see reference 207 for a discussion). Cell lysates can also be fractionated before gel electrophoresis to increase the sensitivity of the method for detecting interaction with rare proteins.

Considerations in affinity blotting include the biological activity of the proteins on the membrane, the preparation of the protein probe, and the method of detection. Denaturing gels, which are run in the presence of SDS and sulfhydryl reducing agents, will inactivate most proteins and separate subunits of a complex. These denaturants are removed during the blotting procedure, which allows many proteins to recover (or partially recover) activity. However, if biological activity is not recoverable, the proteins can be fractionated by a nondenaturing gel system. This method eliminates the problem of regeneration of activity and allows the detection of binding in cases when binding requires the presence of a protein complex.

The protein probe can be prepared by any one of several procedures, and, as with affinity columns, the recent advent of fusing tags to the protein has greatly facilitated this purification. Synthesis in *E. coli* with a GST fusion, epitope tag, or other affinity tag is most commonly used. The protein of interest can then be radioactively labeled, biotinylated, or used in the blotting procedure as an unlabeled probe that is detected by a specific antibody. Vectors that incorporate into the protein a short amino acid sequence recognized by the heart muscle cyclic AMP (cAMP)-dependent protein kinase provide another convenient means for in vitro labeling (18).

One example of affinity blotting is the study of calmodulin-binding proteins (77). Calmodulin can be ^{125}I labeled and used either to probe a gel strip directly or to probe a nitrocellulose membrane after transfer of fractionated proteins. Because the extent of renaturation of calmodulin-binding proteins is variable, the assay is not quantitative. False-positive results have been detected in which a basic sequence binds calmodulin, although generally this binding is Ca^{2+} independent. A major advantage of this technique is that in the analysis of a multimeric protein that binds calmodulin, the precise binding polypeptide can be readily identified by affinity blotting with calmodulin. Using a combination of genetic approaches, Geiser et al. (73) identified the spindle pole body component Spc110 (Nuf1) as interacting with yeast calmodulin and then used affinity blotting to demonstrate that labeled calmodulin could directly detect a GST-Spc110 fusion transferred to a blot after fractionation by SDS-PAGE.

Affinity blotting has been widely used in studies of the association of the regulatory subunit of the type II cAMP-dependent protein kinase with numerous specific anchoring proteins (reviewed in reference 29). Two-dimensional procedures of isoelectric focusing followed by SDS-PAGE have been used to increase the separation of these anchoring proteins. As a control in some of these experiments, a mutant of the regulatory subunit that is deleted for the first 23 residues did not detect any anchoring proteins.

Immunoprecipitation

Coimmunoprecipitation is a classical method of detecting protein-protein interactions and has been used in literally thousands of experiments. The basic experiment is simple. Cell lysates are generated, antibody is added, the antigen is precipitated and washed, and bound proteins are eluted and analyzed. Several sources of material are in wide use. The antigen used to make the antibody can be purified protein (either from the natural tissue or organism or purified after expression in another organism) or synthetic peptide coupled to carrier, and the antibody can be polyclonal or monoclonal. Alternatively, the protein can carry an epitope tag for which commercially available antibodies are available (12CA5 and c-Myc are in common use) or a protein tag (such as GST) for which beads are available to rapidly purify the GST fusion protein and any copurifying proteins. Glutathione-agarose beads were used, for example, to detect and characterize a GTP-dependent interaction of Ras and Raf (227) and to demonstrate that the v-Crk SH2 domain binds the phosphorylated form of paxillin (16). The GST fusion immunoprecipitates a 70-kDa protein that reacts with anti-paxillin antibody and with anti-phosphotyrosine antibody; moreover, anti-paxillin immunoprecipitates a protein that reacts with anti-Crk antibody but only under conditions when the paxillin is phosphorylated.

Several criteria are used to substantiate the authenticity of a coimmunoprecipitation experiment. First, it has to be established that the coprecipitated protein is precipitated by the antibody itself and not by a contaminating antibody in the preparation. This problem is avoided by the use of monoclonal antibodies. Polyclonal antibodies are usually preadsorbed against extracts lacking the protein to remove contaminants or are prepurified with authentic antigen. Peptide-derived antisera (which are usually made by coupling of the peptide to a carrier protein) are usually preadsorbed against the protein that was coupled, to remove antibody against the carrier, in addition to the usual purification to remove contaminating antibody. Second, it has to be established that the antibody does not itself recognize the coprecipitated protein. This can be accomplished by demonstrating persistence of coprecipitation with independently derived antibodies, ideally with specificities toward different parts of the protein. Alternatively, it can sometimes be demonstrated that coprecipitation requires the presence of the antigen; cell lines, growth conditions, or strains that lack the protein cannot coprecipitate the protein unless the antigen is added. In certain cases, it can also be shown that antibody generated against the coprecipitated protein will coprecipitate the original antigen. Third, one would like to determine if the interaction is direct or proceds through another protein that contacts both the antigen and the coprecipitated protein. This is usually addressed with purified proteins, by immunological or other techniques. Fourth, and most difficult, is determining that the interaction takes place in the cell and not as a consequence of cell lysis. Such proteins ought to colocalize, or mutants ought to affect the same process.

A particularly good example of this technique is the demonstration that adenovirus E1A protein interacts with Rb protein. A mixture of monoclonal antibodies against E1A coimmunoprecipitated a discrete set of five polypeptides (and some smaller ones) from a cell line expressing E1A, including a particularly abundant one of 110 kDa (84). Four lines of evidence supported the claim that the 110-kDa polypeptide was forming a complex with E1A protein. First, coprecipitation was not specific to a single antibody; three independent monoclonal antibodies against E1A protein coimmunoprecipitated this protein. Second, these antibodies did not themselves recognize or immunoprecipitate the native or denatured 110-kDa protein, although they recognized and immunoprecipitated native and denatured E1A protein. Third, coprecipitation required E1A protein; the 110-kDa polypeptide could be immunoprecipitated from HeLa extracts (which do not contain E1A protein) only if a source of E1A protein was added. Fourth, the complex could be detected independently in crude lysates; a subpopulation of E1A protein in lysed cells sedimented at 10S rather than at 4S, and this subpopulation contained coimmu-

noprecipitable 110-kDa protein. A similar 110-kDa protein (as well as a similar set of other proteins) was also identified with antipeptide antisera against E1A protein (242). Two separate antisera (one against an amino-terminal peptide and one against a carboxyl-terminal peptide) each coprecipitated the 110-kDa polypeptide, and coprecipitation was prevented either with an excess of the corresponding E1A peptide antigen or in cell extracts lacking E1A protein.

Subsequent studies established that this 105- to 110-kDa polypeptide was the Rb gene product (236). To this end, monoclonal antibodies against the 110-kDa protein were prepared by immune purification of the 110-kDa protein. The resulting antibody coprecipitated E1A protein, just as anti-E1A coprecipitated the 110-kDa protein. Since the 110-kDa protein was the same size as Rb protein, and since it was present in a wide variety of cell lines but not in cell lines known to contain deletions of the Rb gene, it seemed likely that the 110-kDa protein was Rb protein. This was proved by using anti-Rb peptide antibodies against different regions of Rb in three experiments. First, 110-kDa protein precipitated with anti-110-kDa antibody comigrated and had the same partial peptide map as that precipitated with anti-Rb antibody. Second, 110-kDa protein precipitated with anti-E1A antibody could be detected in immunoblots with two different anti-Rb antibodies, and this detection was inhibited by the corresponding peptide antigen. Third, anti-110-kDa antibody could immunoprecipitate Rb protein synthesized in vitro.

When coimmunoprecipitation is performed with unsupplemented crude lysates, as is often the case, this technique has four distinct advantages. First, like protein affinity chromatography, it detects the interactions in the midst of all the competing proteins present in a crude lysate; therefore, the results from this sort of experiment have a built-in specificity control. Second, both the antigen and the interacting proteins are present in the same relative concentrations as found in the cell; therefore, any artificial effects of deliberate overproduction of the test protein are avoided. Third, elaborate complexes are already in their natural state and can be readily coprecipitated; such complexes might otherwise be difficult to assemble in vitro. Fourth, the proteins are present in their natural state of posttranslational modification; therefore, interactions that require phosphorylation (or lack of phosphorylation) are more realistically assessed. Two disadvantages are also apparent. First, coimmunoprecipitating proteins do not necessarily interact directly, since they can be part of larger complexes. For example, the coprecipitation of E1A and p60 (now known to be cyclin A) (84) occurs indirectly; E1A interacts with p107 (237), and p107 interacts with cyclin A (61, 62). Similarly, coprecipitation of Rb protein with E2F probably occurs through another protein (92, 179). Second, coprecipitation is not as sensitive as other methods, such as protein affinity chromatography, because the concentration of the antigen is lower than it is in protein affinity chromatography. This can be overcome by deliberately adding an excess of the antigen to the crude lysates to drive complex formation, as was done to detect a 46-kDa protein that competed with simian virus 40 T antigen for Rb protein (100). It can also be overcome by covalently cross-linking the proteins prior to immunoprecipitation (48) see the section on cross-linking, below). These alterations of course perturb the natural conditions that make immunoprecipitation an attractive method.

Cross-Linking

Cross-linking is used in two ways to deduce protein-protein interactions. First, it is used to deduce the architecture of proteins or assemblies that are readily isolated intact from the cell. Second, it is used to detect proteins that interact with a given test protein ligand by probing extracts, whole cells, or partially purified preparations.

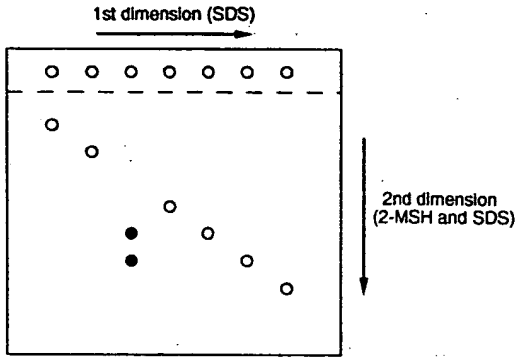


FIG. 2. Two-dimensional gels to identify cross-linked proteins in a complex. Proteins that are not cross-linked have the same mobility in both dimensions of the SDS gel and form a diagonal. Proteins that are cross-linked migrate slowly in the first dimension; after cleavage of the cross-link with mercaptoethanol (2-MSH), these proteins migrate at their native positions in the second dimension and are off the diagonal.

Determination of architecture. The classical method of identifying interacting partners in a purified protein complex involves analysis by two-dimensional gels (Fig. 2). The procedure involves three steps. First, the complex is reacted with a cleavable bifunctional reagent of the form RSSR′, and the R and R′ groups react with susceptible amino acid side chains in the protein complex. This reaction forms adducts of the form P-RSSR′-P′. Second, the proteins are fractionated on an SDS-gel in the absence of reducing agents. The gel separates the proteins based on molecular weight, and cross-linked proteins of the form P-RSSR′-P′ migrate as species of greater molecular weight. Third, a second dimension of the SDS-gel is run after treatment of the gel with a reducing agent to cleave the central S—S bond. Un-cross-linked species align along the diagonal, because their molecular weights do not change after reduction. Cross-linked proteins migrate off the diagonal because they migrated as P-RSSR′-P′ in the first dimension and as molecules of the form P-RSH and P′-R′SH in the second dimension. The cross-links are identified by their size, which corresponds to that of the un-cross-linked species P and P′. This method has been discussed at a practical step-by-step level by Traut et al. (215).

Cross-linking has been used to study the architecture of multienzyme complexes such as CF_1-ATPase (7) and *E. coli* F_1-ATPase (21). It has also been used to study the structure of much more complicated structures like the ribosome (41, 215). Since these structures are complex, the corresponding cross-linking pattern is necessarily complex. Furthermore, as might be expected, different patterns are sometimes obtained as the reactive group is changed and as the distance between the reactive groups is altered (41, 215). Several approaches have been taken to simplify the cross-linking patterns resulting from these experiments. In one approach, the proteins are prefractionated on urea-acrylamide gels or on CM-Sepharose before diagonal electrophoresis (41, 217). A second approach involves running two-dimensional gels without cleaving the cross-link, followed by lution of individual species, cleavage of the cross-link, and resolution of the resulting proteins on a third gel (22). A third approach involves the use of antibody to identify cross-linked partners after the use of appropriate gels (180, 212). Transfer of the gels followed by immunoblotting allows one to

unequivocally identify each member of a cross-linked pair. Since this method is so powerful, one-dimensional gels often suffice and noncleavable cross-linking reagents are easily used. Since immunoblotting is also very sensitive, one can take care to limit cross-linking to acceptably low levels.

Detecti n of interacting proteins. (i) Detection in vivo. Cross-linking in vivo can be accomplished with membrane-permeable cross-linking reagents followed by immunoprecipitation of the ligand protein. This method was used to detect a 60-kDa protein that interacts with Ras (48). Immunoprecipitation of this protein required both immune sera and cross-linking and was inhibited when excess Ras was added before immunoprecipitation. Since the cross-linked protein could be released from the immune complex by cleavage of the cross-link with dithiothreitol (but not by incubation of the immune complex in buffer), it was truly cross-linked. Since pretreatment of the cross-linking reagent with excess amino groups inhibited cross-linking but excess amino groups did not inhibit cross-linking if cells were lysed in their presence, cross-linking must have occurred in vivo. The complex was reproducibly increased after mitogenic stimulation and could be detected in cells producing normal amounts of Ras. This experiment makes another point: at least in these experiments, cross-linking before immunoprecipitation is a more sensitive technique than immunoprecipitation alone.

(ii) Detection in vitro. The addition of an isolated protein or a peptide to a complex system offers a huge potential for precise and powerful cross-linking methods. Several different such methods have been used to detect interacting proteins.

(a) Labeled peptide or protein. Detection of cross-linking partners is incomparably cleaner if the protein or peptide is labeled before cross-linking, because there is only one source of labeled material. For example, ^{125}I-labeled gamma interferon was used to detect receptors that were cross-linked (192), and in vivo labeled interleukin-5 was purified before cross-linking to detect interacting receptors (147).

Proteins are also routinely labeled in vitro with [^{35}S]methionine during translation, and this was followed by cross-linking and by immunoprecipitation to detect protein interactions. This has been done, for example, to detect interaction of preprolactin and pre-β-lactamase with signal sequence receptor and translocation chain-associating protein during glycosylation (79) and to detect mitochondrial import proteins in contact with translocation intermediates (195, 204).

(b) Photoaffinity cross-linking with labeled cross-linking reagent. A particularly useful reagent is the Denny-Jaffee reagent, a cleavable heterobifunctional photoactivatable cross-linking reagent that is labeled on the photoactivated moiety (49). This reagent can be coupled to an isolated protein, which is then incubated in an appropriate extract and photoactivated to cross-link nearby proteins. Since the label is on the photoactivatable moiety of the cross-linking reagent, it is transferred to the cross-linked protein after cleavage of the cross-linking reagent (Fig. 3). This cross-linking reagent has been used to identify a specific 56-kDa ZP3-binding protein on acrosome-intact mouse sperm (19). As much as 90% of the label initially on ZP3 could be transferred to the 56-kDa protein, and cross-linking was inhibited by excess unlabeled ZP3 protein. Moreover, ZP3 affinity columns retained a protein with the same molecular mass. This reagent has also been used to demonstrate that phospholamban interacts with a specific site on the ATPase from sarcoplasmic reticulum only when it is nonphosphorylated and the ATPase is in the Ca^{2+}-free state (106).

Another useful reagent of this type is ^{125}I-{S-[N-(3-iodo-4-azidosalicyl)cysteaminyl]-2-thiopyridine}, also called IAC, a cysteine-specific modifying reagent. This reagent was used to

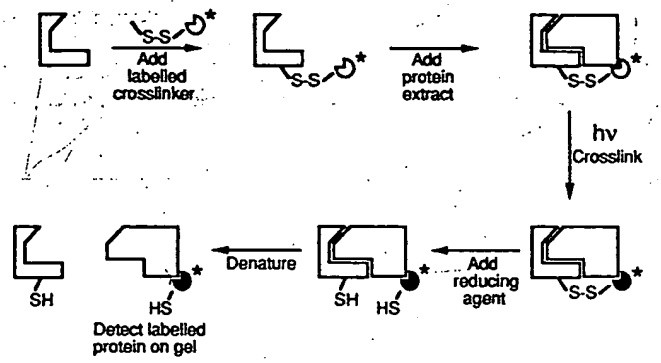


FIG. 3. Specific labeling of an interacting protein with a labeled photoactivatable cross-linking reagent.

demonstrate that the carboxy-terminal region of the subunit of *E. coli* RNA polymerase was adjacent to the activating domain of the catabolite activator protein (CAP) (33). To do this, a unique cysteine was introduced onto the surface of CAP, in a residue which tolerates a large number of mutations, and a preexisting surface cysteine was changed to serine. Subsequent reaction with labeled IAC resulted in quantitative incorporation of label and in protein with 70% of its transcription activation activity. Irradiation of the ternary complex of DNA, CAP, and RNA polymerase yielded 20% cross-linking, all of which was with a particular domain of the subunit of polymerase.

(c) Direct incorporation of photoreactive lysine derivative during translation. A photoactivatable group can be incorporated directly into the translation product by using a modified lysyl-tRNA. If translation is done in the presence of [^{35}S]methionine, the protein is simultaneously labeled and ready for photoactivated cross-linking. This approach has been particularly valuable in investigating the process by which proteins are inserted into the endoplasmic reticulum. During elongation, signal recognition particle (SRP) binds the nascent chain and halts translation until the arrested translation product is brought to the SRP receptor. This releases SRP, allowing translation to continue, coupled with translocation of the protein into the endoplasmic reticulum. With bovine preprolactin, there are two lysines at positions 4 and 9 of the signal sequence and no other lysine residues within the first 70 amino acids, after which translation is normally stopped by SRP. Thus, incorporation of lysine with a photoactivated group specifically probes interaction of the signal sequence with other interacting proteins. In this way, the nascent chain was specifically cross-linked with the 54-kDa protein of SRP and a 35-kDa microsomal membrane protein, called the signal sequence receptor (239). Subsequent experiments in the same system relied on translation of truncated mRNAs bearing lysine codons at different positions. These templates produce proteins that remain tethered to the ribosome through peptidyl-tRNA because of the lack of a termination codon. They therefore cannot complete translocation and are trapped, presumably as intermediates. In this way, it was shown that lysines in different positions also recognized the same 35-kDa membrane protein (121, 238). Moreover, this protein is probably required for translocation because antibodies against it inhibit translocation in vitro (87).

Investigation with the same system in *S. cerevisiae* demonstrated that prepro-α-factor is in contact with Sec61 protein (155). Antibody against either Sec61 or prepro-α-factor precipitated the same labeled cross-linked protein. Cross-linking

was observed only when prepro-α-factor was tethered; release of the protein with puromycin or a complete translation sequence abolished cross-linking. Moreover, the tethered prepro-α-factor was glycosylated while it was tethered, and cross-linking was ATP dependent for large tethered prepro-α-factor peptides; this indicated that prepro-α-factor had entered the normal glycosylation pathway. Sanders et al. (191) also demonstrated by conventional cross-linking followed by immunoprecipitation that Sec61 is in contact with tethered proteins being translocated (in this case by covalent coupling to avidin); the same experiments also demonstrated that BiP (Kar2) was cross-linked to the translocation intermediates and that *sec62* and *sec63* mutants modulate the process. The convergence of genetics and biochemical cross-linking studies further substantiates these interactions.

(d) Site-specific incorporation of photoreactive amino acid derivative during translation. Use of a suppressor tRNA to incorporate a photoactivatable amino acid derivative results in site-specific incorporation by use of a gene carrying a single stop codon. Two such reports have been described. High et al. (94) used a charged amber suppressor tRNA to insert a phenylalanine derivative into various regions of the signal sequence of preprolactin. Cross-linking experiments demonstrated that the amino-terminal end of the signal sequence is in proximity to the translocating chain-associating protein, whereas the hydrophobic core of the sequence contacts Sec61 protein. Cornish et al. (39) used a similar method to incorporate a different photoaffinity label. Still to be described is a similar method involving a labeled photoactivated amino acid replacement—the ultimate magic bullet.

(iii) Other considerations. One major disadvantage of using any cross-linking technique to detect protein-protein interactions is that it detects nearest neighbors which may not be in direct contact. The cross-linking reagent reaches out to any protein in close vicinity; thus, it may appear to detect protein interactions that are more like ships just passing in the night. This is more and more of a problem as the size of the cross-linking reagent increases. Any interaction detected by cross-linking should therefore be independently assessed for protein-protein interactions. However, cross-linking has three important advantages over other methods. First, it can "cement" weak interactions that would otherwise not be visible by other methods (see, for example, reference 48). Second, it can be used to detect transient contacts with different proteins at various stages in a dynamic process such as glycosylation, by freezing the process at different stages. Third, cross-linking can be done in vivo with membrane-permeable cross-linking reagents (48). It may also be possible to detect cross-linking in vivo after microinjection of a protein that is modified with a photoactivatable cross-linking group. To our knowledge, this has not yet been reported.

LIBRARY-BASED METHODS

A variety of methods have been developed to screen large libraries for genes or fragments of genes whose products may interact with a protein of interest. As these methods are by their nature highly qualitative, the interactions identified must be subsequently confirmed by biochemical approaches. However, the enormous advantage of these strategies is that the genes for these newly identified proteins or peptides are immediately available. This is in sharp contrast to the biochemical methods described in the section on physical methods to select and detect proteins that bind another protein, above, which generally result in the appearance of bands on a polyacrylamide gel. These library methods also differ from classical genetic techniques described in the section on genetic methods, below, which often require a specific phenotype before they can be carried out. Library screens are generally performed in bacteria or yeasts, organisms with rapid doubling times. Thus, these procedures can be completed rapidly.

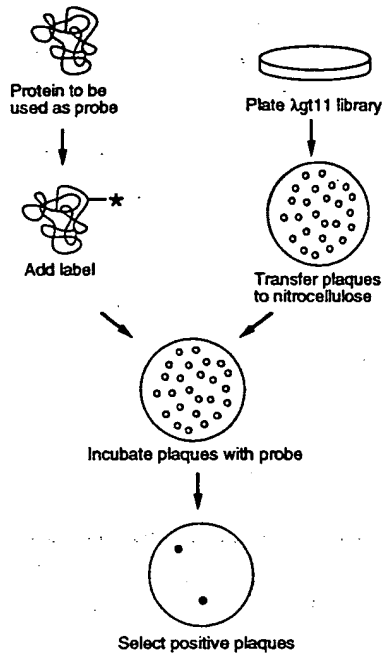


FIG. 4. Use of a labeled protein to probe an expression library.

Protein Probing

A labeled protein can be used as a probe to screen an expression library in order to identify genes encoding proteins that interact with this probe. Interactions occur on nitrocellulose filters between an immobilized protein (generally expressed in *E. coli* from a λgt11 cDNA library) and the labeled probe protein (Fig. 4). The method is highly general and therefore widely applicable, in that proteins as diverse as transcription factors and growth factor receptors have been used as probe. A variety of approaches can be used to label the protein ligand, or this ligand can be unlabeled and subsequently detected by specific antibody.

The method is based on the approach of Young and Davis (244), who showed that an antibody can be used to screen expression libraries to identify a gene encoding a protein antigen. The λgt11 libraries typically use an isopropyl-β-D-thiogalactopyranoside (IPTG)-inducible promoter to express proteins fused to β-galactosidase. Proteins from the bacteriophage plaques are transferred to nitrocellulose filters, incubated with antibody, and washed to remove nonspecifically bound antibody. Protein ligands were first used as probes in this type of experiment by Sikela and Hahn (200), who identified a brain calmodulin-binding protein with ^{125}I-labeled calmodulin as the probe. The λgt11-expressed fusion protein bound calmodulin with a K_d between 3 and 10 nM, and binding was dependent on the presence of Ca^{2+}. The signal-to-noise ratio in these experiments was higher than that found with various antibody probes.

MacGregor et al. (135) used th leucine zipper and DNA-binding domain of Jun as a probe and identified th rat cAMP response element-binding protein type 1. In this case, the Jun

domain was biotinylated and detected with a streptavidin-alkaline phosphatase conjugate. Buffer conditions could be adjusted to distinguish a Jun-Jun homodimer from the more stable Fos-Jun heterodimer. Blackwood and Eisenman (17) used a similar approach with the basic-region helix-loop-helix leucine zipper domain (bHLH-zip) of the c-Myc protein. A 92-residue carboxy terminus of Myc, containing this domain, was expressed as a GST fusion protein, purified by glutathione-agarose affinity chromatography and ^{125}I labeled. This probe identified a new bHLH-zip protein termed Max, and gel shift experiments indicated that the Myc-Max complex exhibited site-specific DNA binding under conditions where neither Myc nor Max alone could bind. These results were critical in establishing a long-sought role for the Myc protein. Extending this result, Ayer et al. (6) used Max as a labeled probe to identify another member of this class, termed Mad.

A major advantage of the protein-probing approach is that the protein probe can be manipulated in vitro to provide, for example, a specific posttranslational modification or a metal cofactor. This modification or cofactor may be essential for the ability of the probe to bind to other proteins. This feature of the approach was exploited in the Ca^{2+}-dependent binding of calmodulin (200). Skolnik et al. (201) extended this use to phosphorylated probes in order to find proteins that bind to the carboxy-terminal phosphorylated tail of the epidermal growth factor (EGF) receptor. This tail is part of the intracellular domain of the receptor, which possesses a protein tyrosine kinase activity stimulated by binding of EGF. Skolnik et al. purified this domain from cells infected with a recombinant baculovirus, tyrosine phosphorylated it in vitro, and cleaved it to separate the phosphorylated carboxy-terminal tail from the protein kinase domain. Probing an expression library identified proteins containing the SH2 domain, which recognizes phosphotyrosyl-containing peptides. This cloning approach might be applied to the identification of proteins interacting with other activated phosphorylated receptors, including tyrosine- and serine-specific phosphatases as well as kinases. In addition, it should be possible to modify probe proteins by means other than phosphorylation to identify new proteins that recognize such modifications.

Probing expression libraries with labeled protein has numerous advantages. Since any protein or protein domain can be specifically labeled for use as a probe, the sophisticated arsenal of GST fusion vectors, other expression and tagging systems, and in vitro translation systems can be exploited; this makes preparation of the probe relatively straightforward. If specific antibody to the target protein is available, the probe protein need not be labeled; the antibody can be used in a second step to detect plaques that have bound the target protein. More than 10^6 plaques can be screened in an experiment, plating 5×10^4 plaques per 150-mm dish. The method not only results in the immediate availability of the cloned gene for the interacting protein but also can provide data regarding a specific domain involved in the interaction, because the λgt11 insert is often only a partial cDNA. Conditions of the wash cycles can be adjusted to vary the affinity required to yield a signal. As with many library-based methods, probing expression libraries compares equally all binary combinations of the probe protein and a library-encoded protein. Thus, less abundant proteins, proteins with weak binding constants, and proteins that temporally or spatially rarely interact with the probe protein in vivo can all be detected as long as their transcripts are present in the mRNA pool used to generate the library.

This method has certain intrinsic limitations. Proteins encoded by the library must be capable of folding correctly in *E. coli*, generally as fusion proteins, and of maintaining their structure on a nitrocellulose filter. However, proteins often can be renatured by subjecting the filters to a denaturation-renaturation cycle with 6 M guanidine hydrochloride as described by Vinson et al. (222). Binding conditions are arbitrarily imposed by the investigator, rather than reflecting the native environment of the cell. Since all combinations of protein-protein interactions are assayed, including those that might never occur in vivo, the possibility of identifying artifactual partners exists. In particular, the relative abundance of each potential partner expressed in a colony or plaque of the library is similar, instead of varying and potentially being compartmentalized as in the cell. Any posttranslational modifications necessary for efficient binding will generally not occur in bacteria (although some such modifications can be performed in vitro). Screening rather than direct selection is the means of detection, which inherently limits the number of plaques that can be assayed. The use of screening also restricts the further genetic manipulations that can be applied to the cDNA inserts. For example, in the analysis of point mutations, it is not possible to select directly for rare mutations that affect the interaction. Different protein probes are likely to behave variably in this approach, such that binding and washing conditions may have to be adjusted in each case to maximize the signal-to-noise ratio.

Phage Display

Basic approach. Smith (203) first demonstrated that an *E. coli* filamentous phage can express a fusion protein bearing a foreign peptide on its surface. These foreign amino acids were accessible to antibody, such that the "fusion phage" could be enriched over ordinary phage by immunoaffinity purification. Smith suggested that libraries of fusion phage might be constructed and screened to identify proteins that bind to a specific antibody. In the past few years, there have been numerous developments in this technology to make it applicable to a variety of protein-protein and protein-peptide interactions.

Filamentous phages such as M13, fd, and f1 have approximately five copies of the gene III coat protein on their surface; thus, a foreign DNA sequence inserted into this gene results in multiple copies of the fusion protein displayed by the phage. This is called polyvalent display. Similarly, the major coat protein encoded by gene VIII can also display a foreign insert (104). The gene VIII protein allows up to 2,700 copies of the insert per phage. Generally, polyvalent display is limited to small peptides (see the next section) because larger inserts interfere with the function of the coat proteins and the phage become poorly infective.

Random sequences can be inserted into gene III or gene VIII to generate a library of fusion phage (Fig. 5). Such a library can then be screened to identify specific phage that display any sequence for which there is a binding partner, such as an antibody. This screening is performed by a series of affinity purifications known as panning. The phage are bound to the antibody, which is immobilized on a plastic dish. Phage that do not bind are washed away, and bound phage are eluted and used to infect *E. coli*. Each cycle results in a 1,000-fold or greater enrichment of specific phage, such that after a few rounds, DNA sequencing of the tight-binding phage reveals only a small number of sequences. In addition to the advantage of high selectivity, a second advantage of this technology is that large phage libraries can be constructed (up to 10^9 to 10^{10} complexity) and the affinity purification step can be carried out at very high concentrations of phage ($>10^{13}$ phage per ml) (50). Third, the direct coupling of the fusion protein to its gene in a single phage allows the immediate availability of sequence

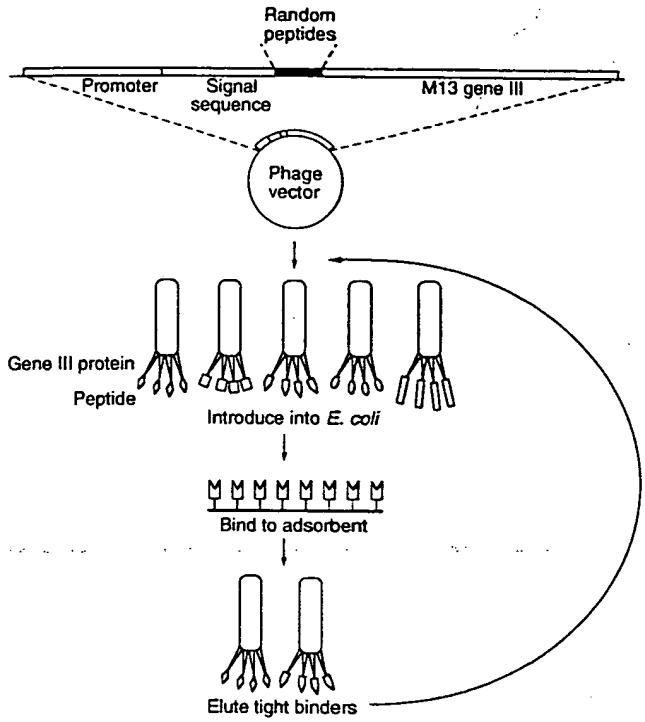


FIG. 5. A peptide library in a filamentous phage vector. The figure illustrates the process of panning, by which peptides that bind to an adsorbent are identified.

data to generate one or more consensus sequences of bound peptides or the sequences of variant proteins with a specific phenotype. Fourth, the phage can be used directly to assess the binding specificity of the encoded fusion proteins by varying the stringency of the wash procedures used in the panning cycles.

Random-sequence peptide libraries have been generated by cloning synthetic oligonucleotides into gene III (Fig. 5). Scott and Smith (198) generated a hexapeptide library and screened it to identify epitopes for two monoclonal antibodies specific for a hexapeptide from the protein myohemerythrin. Cwirla et al. (44) constructed a similar hexapeptide library to find peptides that can bind to a monoclonal antibody specific for a tetrapeptide from β-endorphin. Such epitope libraries allow rapid characterization of an unknown epitope recognized by either a monoclonal antibody or polyclonal serum. For example, monoclonal antibody pAB240, which recognizes the mutant conformation of the tumor suppressor p53 protein, was shown to bind to a 5-amino-acid motif in p53 (210). The binding partner for the phage-encoded peptides need not be an antibody. For example, Devlin et al. (50) constructed a 15-residue peptide library and used it to identify nine different peptides that bind to streptavidin.

A major advance in phage display came with the development of a monovalent system in which the coat protein fusion is expressed from a phagemid and a helper phage supplies a large excess of the wild-type coat protein (11, 131). Therefore, the phage are functional because the recombinant protein makes up only a small amount of the total coat protein. The vast majority (>99%) of the population of phage particles display either one or no copies of the fusion protein on their surface. Such phage can accommodate 50 kDa of foreign protein without any significant effect on phage infectivity. In addition, monovalent phage display avoids potential avidity effects observed with polyvalent display, in which the phage can attach to the adsorbent at multiple points.

Phage display has also been used to identify proteins with increased binding affinity. In some cases, the use of monovalent display was necessary to avoid potential avidity effects, attributed to multipoint attachment of the polyvalent phage to the absorbent (231). Lowman et al. (131) expressed nearly one million mutants of human growth hormone (191 residues) as fusion phage and identified variants that bound tightly to the growth hormone receptor. The mutations were directed to 12 sites known to be important for binding to the receptor. Some variants had binding affinities up to eightfold greater than that of the wild-type hormone. Roberts et al. (186) used polyvalent display of bovine pancreatic trypsin inhibitor and directed mutagenesis to five residues of the protein. They selected for high-affinity inhibitors of human neutrophil elastase and identified one variant with an affinity 3.6×10^6 higher than that of wild-type bovine pancreatic trypsin inhibitor.

A similar strategy can be used with nontargeted mutagenesis. For example, Pannekoek et al. (167) expressed human plasminogen activator inhibitor 1, a 42-kDa protein, as a gene III protein fusion under conditions for monovalent display. The phage-displayed inhibitor could specifically form complexes with serine protease tissue-type plasminogen activator. PCR mutagenesis was used to generate a library of mutant plasminogen activator inhibitor 1 proteins, which can be screened to analyze structure-function relationships.

Phage display presents several advantages for the study of protein-protein interactions. The very large sizes of either random libraries or pools of individual variants of a single sequence that can be generated mean that complex mixtures can be screened. While not strictly a genetic approach, in that there is no direct selection for an interacting partner, phage display has many of the properties of genetic selection through its use of panning cycles. It is a rapid procedure and should be widely applicable. Although screening a random library of cDNA by a panning procedure to identify proteins that interact with a protein of interest has not yet been demonstrated, this strategy should prove workable.

Disadvantages of phage display include the size limitation of protein sequence for polyvalent display; the requirement for proteins to be secreted from *E. coli*; and the use of a bacterial host which may preclude the correct folding or modification of some proteins. All phage-encoded proteins are fusion proteins, which may limit the activity or accessibility for binding of some proteins. Since binding is detected in vitro, the same considerations of an in vitro approach that are relevant for protein probing of expression libraries are relevant here.

Related methods. (i) Antibody phage. While we do not specifically address the vast topic of antigen-antibody interactions in this review, it is worth noting that phage display can be applied to these interactions. The principle of displaying antibody-combining domains on the surface of phage was first demonstrated by McCafferty et al. (141). The heavy- and light-chain variable domains of an anti-lysozyme antibody were linked on the same polypeptide and expressed as a gene III protein fusion. Over 1,000-fold enrichment of the antibody could be obtained by a single passage over a lysozyme-Sepharose column. This method was then extended by this and other groups to allow the display of libraries of combining domains, such that new antibodies or mutant versions of existing antibodies could be generated.

Kang et al. (110) used a vector to express a combinatorial library of functional Fab molecules (~50-kDa heterodimer) on the surface of a phage. The Fd chain, consisting of the variable region and constant domain 1 of the immunoglobulin heavy

chain, was synthesized as a gene VIII protein fusion, while the κ light chain contained no phage sequence. The two chains could assemble in the bacterial periplasm and become incorporated into the phage on coinfection with helper. Phage contained 1 to 24 antigen-binding sites per particle. The vector system described allows recombination of the two chains to generate large combinatorial libraries. A similar strategy to express Fabs by using the gene III protein has also been described (10). Additionally, a combinatorial library of linked heavy- and light-chain variable genes fused to the gene III protein has been shown to be capable of detecting a high-affinity binder (37). Kang et al. (110) suggested that such systems can be used for mutation and selection cycles to generate high-affinity antibodies. Moreover, they envisioned that the systems can be extended to analyze any protein recognition system, such as ligand-receptor interactions.

Phage display of Fab fragments was extended by Burton et al. (26), who generated a library of such fragments from the RNA of a human immunodeficiency virus-positive individual. After four rounds of panning with immobilized surface glycoprotein gp120 of the virus as the adsorbent, specific viral antibodies were obtained. A similar method was used to obtain human antibody Fabs that recognize the hepatitis B surface antigen (246).

(ii) Peptides on plasmids. In a method highly analogous to phage display, random peptides are fused to the C terminus of the *E. coli* Lac repressor and expressed from a plasmid that also contains Lac repressor-binding sites (43). Thus, the peptide fusions bind to the same plasmid that encodes them. The bacterial cells are lysed, and the peptide libraries are screened for peptides that bind to an immobilized receptor by using similar panning cycles to those for phage libraries. In this case, peptides become enriched because bound peptides carry their encoding plasmids with them, via the repressor-operator interaction, and these plasmids can be transformed back into *E. coli*. In the initial example, peptides that bind to a monoclonal antibody specific for dynorphin B were selected, and these peptides contained a hexapeptide sequence similar to a segment of dynorphin B (43). This method is distinguished from the phage display methods in that the peptides are exposed at the C terminus of the fusion protein and the fusions are cytoplasmic rather than exported to the periplasm.

Two-Hybrid System

The two-hybrid system (35, 65, 66) is a genetic method that uses transcriptional activity as a measure of protein-protein interaction. It relies on the modular nature of many site-specific transcriptional activators, which consist of a DNA-binding domain and a transcriptional activation domain (23, 97, 112). The DNA-binding domain serves to target the activator to the specific genes that will be expressed, and the activation domain contacts other proteins of the transcriptional machinery to enable transcription to occur. The two-hybrid system is based on the observation that the two domains of the activator need not be covalently linked and can be brought together by the interaction of any two proteins. The application of this system requires that two hybrids be constructed: a DNA-binding domain fused to some protein, X, and a transcription activation domain fused to some protein, Y. These two hybrids are expressed in a cell containing one or more reporter genes. If the X and Y proteins interact, they create a functional activator by bringing the activation domain into close proximity with the DNA-binding domain; this can be detected by expression of the reporter genes (Fig. 6). While the assay has been generally performed in yeast cells, it works similarly in mammalian cells

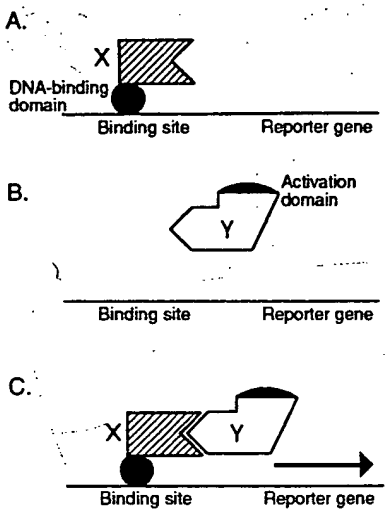


FIG. 6. The two-hybrid system. (A) The DNA-binding domain hybrid does not activate transcription if protein X does not contain an activation domain. (B) The activation domain hybrid does not activate transcription because it does not localize to the DNA-binding site. (C) Interaction between X and Y brings the activation domain into close proximity to the DNA-binding site and results in transcription.

(see, e.g., reference 46) and should be applicable to any other eukaryotic cells.

This method has been used with a wide variety of proteins, including some that normally reside in the nucleus, cytoplasm, or mitochondria, are peripherally associated with membranes, or are extracellular (see reference 66 for a review). It can be used to detect interactions between candidate proteins whose genes are available by constructing the appropriate hybrids and testing for reporter gene activity (220, 249). If an interaction is detected, deletions can be made in the DNA encoding one of the interacting proteins to identify a minimal domain for interaction (35). In addition, point mutations can be assayed to identify specific amino acid residues critical for the interaction (127). Most significantly, the two-hybrid system can be used to screen libraries of activation domain hybrids to identify proteins that bind to a protein of interest. These screens result in the immediate availability of the cloned gene for any new protein identified. In addition, since multiple clones that encode overlapping regions of protein are often identified, the minimal domain for interaction may be readily apparent from the initial screen (105, 223).

A variety of versions of the two-hybrid system exist, commonly involving DNA-binding domains that derive from the yeast Gal4 protein (35, 55) or the *E. coli* LexA protein (223, 247). Transcriptional activation domains are commonly derived from the Gal4 protein (35, 55) or the herpes simplex virus VP16 protein (45). Reporter genes include the *E. coli lacZ* gene (65) and selectable yeast genes such as *HIS3* (55) and *LEU2* (247). An increasing number of activation domain libraries are becoming available, such that screens are now feasible for proteins from many different organisms or specific mammalian tissues.

One field in which the two-hybrid system has been applied with considerable success has been the study of oncogenes and tumor suppressors and the related area of cell cycle control. For example, reconstruction experiments with previously cloned proteins indicated that interactions occur between Ras and the protein kinase Raf (220, 249), human Sos1 guanine nucleotide exchanger and the growth factor receptor-associ-

ated protein Grb2 (30), and Raf and the transcription factor inhibitor IκB (129). Two-hybrid searches with oncoproteins or tumor suppressors as targets have identified a leucine zipper protein that binds to Jun (34); protein phosphatase PP1α2, which binds to Rb (55); a bHLH-zip protein Mxi1, which binds to the Myc-associated protein Max (247); and the Rb-related protein p130, which binds to cyclins and was identified through its interaction with the cyclin-dependent kinase Cdk2 (83). A notable convergence of different approaches came about with the identification of another protein that binds to Cdk2, a 21-kDa protein termed Cip1, which inhibits the kinase activity (85). This protein turned out to be identical to a protein encoded by the major p53-inducible transcript (58), suggesting that the tumor suppressor role of p53 may be mediated by its activation of the gene for this 21-kDa protein.

The two-hybrid system has several features that make it useful for analysis of protein-protein interactions. It is highly sensitive, detecting interactions that are not detected by other methods (see, e.g., references 127 and 220). On the basis of binding of different proteins to the retinoblastoma protein, Durfee et al. (56) estimate that the minimal binding constant required to detect an interaction in their version of the two-hybrid system is on the order of 1 μM. This value suggests that the system should be applicable to a wide range of protein interactions. However, it is clear that the minimal affinity interaction detectable will depend on such variables as the level of expression of the hybrid proteins; the number, sequence, and arrangement of the DNA-binding sites in the reporter gene(s); and the amount of reporter protein required for a detectable phenotype. Given these variables, it is likely that some versions of the system may detect weak interactions with binding constants considerably greater than 1 μM. Another advantage is that the interactions are detected within the native environment of the cell and hence that no biochemical purification is required. The use of genetic-based organisms like yeast cells as the hosts for studying interactions allows both a direct selection for interacting proteins and the screening of a large number of variants to detect those that might interact either more or less strongly. With a reporter gene such as the yeast *HIS3* gene, the competitive inhibitor 3-aminotriazole can be used to directly select for constructs which yield increased affinity.

The two-hybrid system is limited to proteins that can be localized to the nucleus, which may prevent its use with certain extracellular proteins. Proteins must be able to fold and exist stably in yeast cells and to retain activity as fusion proteins. The use of protein fusions also means that the site of interaction may be occluded by one of the transcription factor domains. Interactions dependent on a posttranslational modification that does not occur in yeast cells will not be detected. Many proteins, including those not normally involved in transcription, will activate transcription when fused to a DNA-binding domain (134), and this activation prevents a library screen from being performed. However, it is often possible to delete a small region of a protein that activates transcription and hence to remove the activation function while retaining other properties of the protein.

Other Library-Based Methods

A number of other library strategies have been developed recently that, in principle, should result in the identification of proteins that interact with a protein of interest. However, because the first description of methods generally involves known combinations of proteins, the general applicability of a new method cannot be easily judged.

In one approach, the ability of the *E. coli* bacteriophage λ repressor to dimerize was used as a reporter for the interaction of leucine zipper domains (98). The N-terminal domain of repressor binds to DNA but dimerizes inefficiently; a separate C-terminal domain that mediates dimerization is required for efficient binding of the protein to its operator. The N-terminal DNA-binding domain was fused to the leucine zipper of the yeast Gcn4 protein, which allowed dimerization and repression of transcription in *E. coli*. This repression enabled the host cell to survive superinfection by λ phage. This phenomenon enabled Hu et al. (98) to introduce single-amino-acid mutations into the leucine zipper domain and to use a genetic assay in *E. coli* to determine whether dimerization of the zipper domain occurred. They suggested that this assay could be used to select clones from a library for proteins that bind to a target protein, which is expressed in *E. coli* as a repressor hybrid. Any phage that express a protein that binds to the target protein should compete for dimerization of the repressor and its ability to bind λ operators. These phage would be detected because they result in plaques. As described, this approach would be limited to target proteins that homodimerize. In addition, this method when applied to library screening is a competition assay; it would require that the library-encoded protein bind to the target protein in preference to the target protein interacting with itself.

Another *E. coli*-based assay involves tagging the target protein with biotin by fusing it to the biotin carboxylase carrier protein (74). This tag allows the protein to be bound by avidin, streptavidin, or anti-biotin antibody-coated filters. Potential interacting proteins are fused to the LacZ protein and expressed from a λ vector such that β-galactosidase activity is intact. These phage are infected into cells containing the biotin-tagged target protein, and interaction can occur in vivo between a library-encoded protein and the target protein. This interaction is then detected when the phage plaques are transferred to avidin filters and assayed for β-galactosidase activity. The method was shown to work by using biotinylated c-Jun protein and a c-Fos–LacZ fusion. Although the protein-protein interaction occurs within the living bacterial cells, the detection of this interaction occurs in vitro on filters that must be washed after transfer of the proteins. Thus, in principle, this method may have many of the same limitations that protein probing of expression libraries has.

GENETIC METHODS

For organisms for which powerful genetic analysis methods exist, sophisticated strategies can be designed to uncover genes that show interactions with other genes. In many cases, these newly uncovered genes encode proteins that physically interact with proteins encoded by the known genes. In other cases, genetic methods can be used to confirm interactions among previously identified proteins. These strategies are generally based on classical genetic approaches. For example, identification of extragenic suppressors often reveals mutations in genes whose products physically interact with the protein containing the original defect. Synthetic lethal screens yield mutations that, in combination with another nonlethal mutation, result in the inability of the organism to grow; this phenotype is commonly due to alterations in interacting proteins. Overproduction of certain proteins can lead to the suppression of mutations in interacting proteins. In other cases, overproduction disrupts a cellular process by altering the balance of the different components of a complex structure, or th overproduced protein is nonfunctional and acts in a dominant-negative manner.

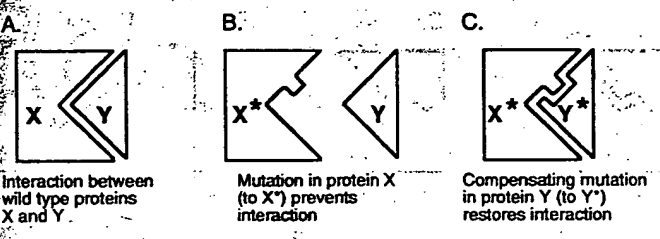


FIG. 7. Extragenic suppression due to restoration of a protein-protein interaction.

The value of some of these genetic approaches has been significantly increased by applying them to organisms not amenable to classical genetic techniques, using modern molecular tools. For example, the ability to generate mice either carrying novel genetic information or deleted for one or more of their endogenous genes allows this organism to be analyzed by some of the logic formerly reserved for much simpler creatures. However, it must be kept in mind with any genetic approach that identification of mutants with the correct phenotypes does not guarantee that the biochemical mechanisms invoked to explain these phenotypes are correct.

Extragenic Suppressors

Suppressor mutations are mutations that partially or fully revert the phenotype caused by an original mutation (see reference 86 for review). Extragenic suppressors occur in genes other than the gene carrying the primary mutation. This is illustrated in Fig. 7, in which a mutation of protein Y to Y* compensates for the defect X* to restore activity to the XY dimer. However, analysis of these suppressors is often difficult, because they lack any phenotype in the absence of the primary mutation. To circumvent this problem, Jarvik and Botstein (107) sought suppressors of temperature-sensitive mutants of phage P22 that resulted in a cold-sensitive phenotype. This cold-sensitive phenotype did not necessarily depend upon the presence of the original mutation causing temperature sensitivity, and thus mutations in new genes could be uncovered. It was proposed (107) that one mechanism of this suppression is that the original mutation and the suppressor lie in genes whose products physically interact and that the original mutation destroyed this interaction. The suppressor then produces a compensating alteration that restores the interaction.

This type of suppressor analysis has been exploited in studying fundamental processes in yeast cells, particularly cell cycle control, cytoskeleton structure, and RNA splicing. Moir et al. (152) isolated cold-sensitive cell division cycle (*cdc*) mutants of *Saccharomyces cerevisiae* and used them to identify temperature-sensitive revertants. Some of these revertants carried new mutations that alone resulted in a *cdc* phenotype at the restrictive temperature, suggesting that the mutated gene products might interact with the cold-sensitive protein. These results support the idea that only a few genes might be capable of mutation to generate an altered product that can suppress the original mutation. Thus, this approach can be applied to a process such as cell cycle control and reveal most or all of the interacting gene products.

In a similar strategy, suppressors of a temperature-sensitive mutation in the *S. cerevisiae* actin gene that acquired a cold-sensitive phenotype identified five new genes (160). Mutations in these genes, even in a background with the wild-type actin gene, led to phenotypes similar to those of actin mutants. These results suggested that these genes could encode proteins that are part of the actin cytoskeleton. In a related approach, dominant suppressors of an actin mutation also identified a gene whose product may interact with actin (3). In both these cases, the suppressor mutations showed allele specificity; some but not all actin alleles were suppressed by a given mutation. This allele specificity also supported the idea of a direct physical interaction, in that suppressor mutations that simply bypass the requirement for the protein containing the original mutation would not be expected to show such specificity.

The nematode *Caenorhabditis elegans* has also been used extensively for suppression analysis because large populations of individuals can be examined (96). If a temperature-sensitive mutant is available, it can be shifted to the restrictive temperature to apply a direct selection for suppressors. This approach has been used to study such processes as movement, egg laying, and sex determination. One example is the suppression of an *unc-22* mutation that resulted in muscle twitching (151). Some of these suppressors were mutations in the *unc-54* gene which encodes the major myosin gene. These results suggested that the *unc-22* and *unc-54* proteins physically interact, and this idea is supported by the finding that the *unc-22* protein, like myosin, is located in the A-bands of muscle (150).

Suppressor analysis can clearly uncover new mutations that affect a process under study, and analysis of the genes and proteins defined by these mutations sometimes indicates interacting proteins. While often used with temperature-sensitive and cold-sensitive mutations, many other types of spontaneous mutations can also be readily suppressed if an appropriate genetic selection is available. With the availability of numerous cloned genes, conditional alleles can now be generated by in vitro mutagenesis methods. An obvious limitation of this type of analysis is that it can generally be applied only to simple organisms such as phages, bacteria, yeasts, nematodes, and *Drosophila* species. It requires not only the gene of interest but also a useful mutant to initiate the analysis. For example, suppressors in an interacting protein may be difficult or impossible to obtain if the original mutation does not affect a domain of interaction. Furthermore, other mechanisms can yield suppressors. These include second intragenic mutations, gene duplication of the original mutant gene, suppression by epistasis, and informational suppression (see, for example, reference 96). Thus, identification of the suppressors of interest against a background of these other mutations can be a time-consuming process.

Synthetic Lethal Effects

Mutations in two genes can cause death (or another observable defect) while mutation in either alone does not. This phenomenon is called a synthetic effect and can result from physical interactions between two proteins required for the same essential function. This is illustrated in Fig. 8, in which the dimer XY is required for some function and loss of this function results in a detectable phenotype. Mutation in X or Y yields partial binding, but the double mutant X*Y* has no binding. Dobzhansky (52) first described synthetic lethal effects in *Drosophila* species. However, the search for synthetic lethal effects has been applied successfully most often in *S. cerevisiae*. One of the tools available for research in this organism is a colony-sectoring assay (93, 119), in which cells containing a plasmid are red and can therefore be easily distinguished from those that have lost the plasmid and are white. If maintenance of the plasmid is not essential for viability of the yeast, colonies appear with red-and-white sectoring. If the cells become dependent on a gene carried by the plasmid, the colonies appear uniformly red. For example, Bender and Pringle (15) used such an assay with a plasmid-borne copy of the *MSB1* gene, which

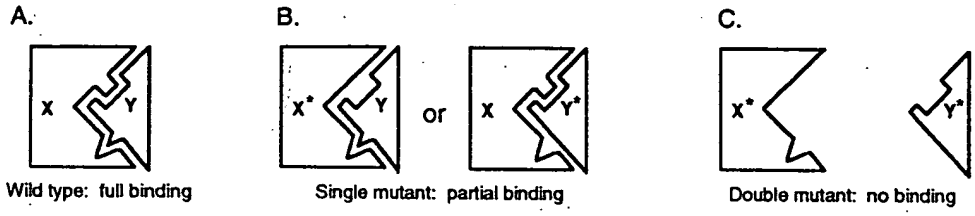


FIG. 8. Synthetic effect, in which either single mutant is functional but not the double mutant.

plays a role in bud formation. They mutagenized the plasmid-containing cells and screened for mutants in which *MSB1* had become essential for survival. This screen identified two new genes, *BEM1* and *BEM2*, in which mutations led to defects in cell polarity and bud emergence. In this approach, if the plasmid is maintained at high copy number in *S. cerevisiae*, it is also possible to identify mutations in new genes that are lethal but can be suppressed by multiple copies of the plasmid-borne gene.

A similar approach was taken by Costigan et al. (40) to identify mutants that require the Spa2 protein, which is also involved in polarized cell growth as well as in the morphogenetic changes that occur in yeast mating. The synthetic lethal screen identified the *SLK1* gene, which is necessary for morphogenesis in vegetatively growing yeast cells and in mating pheromone-treated cells. Costigan et al. pointed out that the synthetic lethal screen by the colony color assay is extremely sensitive and can identify mutants with low viability. Since both *spa2* and *slk1* mutants are individually healthy, the screen did not simply combine two mutations each causing unhealthiness to result in death, a common concern in using this method. Instead, it seems likely that the synthetic lethal effect often results from two different defects in the same cellular process.

Other synthetic lethal screens in yeast cells involve a poison assay in which the presence of a plasmid-borne gene on a particular medium is lethal; when yeast cells containing this plasmid are placed on such a medium, there is strong selection for cells that have lost the plasmid. However, mutants that cannot survive without the plasmid can be identified, because the plasmid also contains the gene of interest whose presence is required in these mutants. Such mutants do not grow on replica plates containing the poison. This approach was used to identify mutations in the 3-hydroxy-3-methylglutaryl coenzyme A reductase genes (12). Alternatively, the gene of interest can be expressed by using a regulated promoter, such that mutants that do not survive the repressed condition are identified. Inducible expression of the yeast *RAS2* gene led to the identification of mutations in the *CYR1* gene, which encodes adenylate cyclase (149). Finally, synthetic lethal effects can be uncovered by combining mutations identified in other genetic screens. For example, yeast cells containing a temperature-sensitive mutation in the *SEC4* gene, essential for secretion, are inviable at the permissive temperature when they also contain a temperature-sensitive mutation in certain other *SEC* genes (190). Yeast cells with mutations in both α-tubulin and β-tubulin are inviable (101).

While synthetic lethal screens often lead to the identification of interacting gene products, other explanations do not require this physical interaction (101). For example, the two proteins might both be components of the same structure, or one protein could regulate the activity of the other. Additionally, there are likely to be some cases in which the combination of two mutations, either of which causes poor growth on its own, leads to complete inviability.

Overproduction Phenotypes

Overproduction of wild-type proteins. The overproduction of some wild-type proteins can lead to phenotypes that provide insight into protein-protein interactions. In *S. cerevisiae*, a multicopy plasmid often suppresses mutations in genes other than the one carried on the plasmid (reviewed in reference 182). For example, a temperature-sensitive mutation in the *CDC28* gene, which encodes a protein kinase involved in controlling cell division, can be suppressed by multicopy plasmids carrying the *CLN1* or *CLN2* gene, which encode cyclins (82).

In other cases, overproduction of a protein can cause a phenotype that is altered by overproduction of an interacting protein. High-copy-number plasmids expressing either of the yeast histone pairs H2A and H2B or H3 and H4 caused an increased frequency of chromosome loss (142). However, overproduction of both pairs of histone proteins did not affect the fidelity of chromosome transmission, indicating that it is the imbalance of the two dimer sets with respect to one another that affects this fidelity (142). Overproduction of the yeast Gal4 protein, the transcriptional activator of the galactose-inducible genes, leads to galactose-independent transcription. However, proper regulation is restored if the Gal80 protein, a negative regulator that binds to the Gal4 protein, is also overproduced (159). While the phenotype due to an overproduced wild-type protein may reflect interactions with another protein (either mutant or wild type), there are several other mechanisms by which such phenotypes can occur. For example, an overproduced protein may bypass the transcriptional regulation due to another protein. In other cases, an overproduced protein may lead indirectly to the stabilization of a mutant protein.

Overproduction of mutant proteins. Overproduction of a nonfunctional version of a protein can result in a mutant phenotype due to disruption of the activity of the wild-type protein (Fig. 9) (reviewed in reference 90). The existence of such dominant-negative proteins can lead to a definition of the oligomerization domain of a protein. An early example of this came from studies of the *E. coli* Lac repressor, which has distinct domains for DNA binding and for oligomerization. A mixed oligomer of wild-type subunits and mutant subunits unable to bind DNA results in a nonfunctional repressor (143). This kind of mutant provides evidence for the multimeric nature of the repressor, and analysis of the sites of mutation defined the domains involved in DNA binding and in oligomerization.

A similar mechanism may operate in many human cancers. The wild-type p53 protein is a transcriptional regulator which is tetrameric, and its oligomerization domain is near the C terminus. Mutations in the central domain of p53 that occur in tumors produce dominant-negative mutant proteins that bind to and inactivate the function of the wild-type protein (67). The ability to manipulate cloned genes and reintroduce these mutant versions into cells now allows dominant-negative mutants to be created in many different organisms. For example,

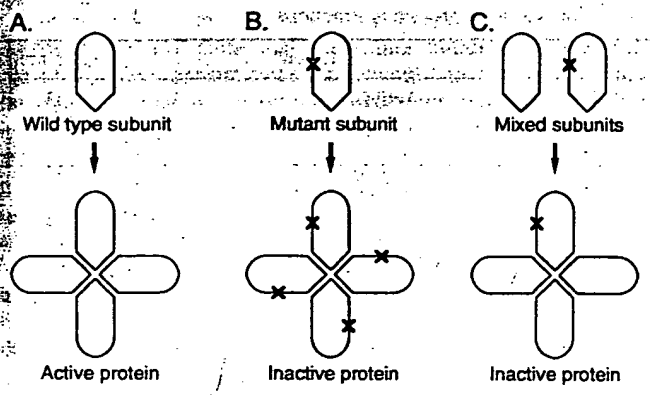


FIG. 9. Dominant-negative effect. Pure populations of wild-type (A) or mutant (B) subunits result in an active or inactive protein, respectively. A mixture of the two types (C) will also be inactive if the mutant subunit acts in a dominant fashion.

dominant-negative Myc proteins were overexpressed in fibroblasts and shown to inhibit transformation by the v-*abl* and *BCR-ABL* oncogenes (194). It was suggested that this effect was due to the mutant Myc proteins competing with the endogenous wild-type Myc protein for binding to the Max protein, thus forming nonfunctional heterodimers.

Unlinked Noncomplementation

Individuals heterozygous for two different recessive mutations sometimes display a mutant phenotype. This unlinked noncomplementation is often interpreted as being due to mutation in two genes that encode interacting products. In *Drosophila* spp., new recessive mutations were identified that failed to complement β_2-tubulin mutations and that mapped to other genes (176). At least one of these mutations mapped very close to an α-tubulin gene. A model for this noncomplementation is based on a minimal dosage requirement for the product of two interacting proteins. If the mutant proteins assemble randomly with the wild type, the double heterozygote would contain only one-fourth the normal level of complex, which would be insufficient for function. In addition, when homozygous, some of the second-site noncomplementing mutations lead to defects in tubulin function, and this property is consistent with the model.

POPULAR METHODS TO ESTIMATE AND DETERMINE BINDING CONSTANTS

Importance of Characterization of the Binding Interaction

The ultimate goal of studying protein-protein interactions is to understand the consequences of the interaction for cell function. This depends in turn on understanding the strength of the interaction in the cell. The determination that two proteins can interact with one another is only the first step in understanding if, and to what extent, the interaction takes place in vivo. Evaluation of the interaction requires the assessment of at least six parameters, which are discussed below.

Binding constant. For any simple interaction of one protein (P) with another (L, for ligand), the interaction is governed by the binding constant K_d, according to the simple equation $K_d = [P_f][L_f]/[PL]$. In this equation, $[P_f]$ and $[L_f]$ refer to the free (i.e., unbound) concentrations of P and L respectively. The interaction between protein and ligand is also expressed in two other ways. First, it is often expressed instead as an affinity constant, $K_a = [PL]/[P_f][L_f]$, i.e., $K_a = 1/K_d$. Second, it is often expressed as a ratio of two rate constants. The rate of formation of PL is $k_a [P_f][L_f]$, where k_a is the association rate constant, and the rate of breakdown of PL is $k_d [PL]$, where k_d is the dissociation rate constant. At equilibrium, the rate of formation of PL equals the rate of breakdown of PL, and $K_d = k_d/k_a$. Evaluation of the dissociation constant is the subject of this section.

Concentrations of species. To evaluate the extent to which two proteins can interact, the cellular (or compartmental) concentrations of P_t (the sum of bound and unbound concentrations) and L_t are required, in addition to the dissociation constant. These two parameters can drastically alter an evaluation of the population of molecules in a complex. For example, if $K_d = [P_t] = [L_t]$, 38% of the species are in the complex PL at any one time. If K_d is 10-fold higher (weaker binding), only 8.4% of the species are in the complex at one time, and if K_d is 10-fold lower (stronger binding), 73% of the species are in a complex. A similar effect holds for alterations in the concentrations of P and L in the cell. A simple way of calculating [PL] from the easily measured parameters $[P_t]$ and $[L_t]$ is as follows: $[PL] = \{([P_t] + [L_t] + K_d)/2\} - 1/2\{([P_t] + [L_t] + K_d)^2 - 4[L_t][P_t]\}^{1/2}$ (54).

Influence of competing proteins. Even if a protein has high affinity for a ligand protein, L, and the protein and ligand are present in sufficient quantities to interact functionally in the cell, they may not do so in vivo to the same extent as in vitro. Other ligands may effectively compete for the ligand protein if they are present at high enough concentration and interact with sufficient affinity. For example, if the concentration of P and L1 are both equal to the dissociation constant, 38% of the species are in a complex. If another ligand, L2 (or a set of potential ligands), is present at 1,000 times the concentration of L1 and has 10-fold-lower affinity for P, the interaction of P with L2 will titrate the vast majority of the protein P (99%, if L2 was the only interacting protein), leaving very little to interact with L1. This sort of consideration is addressed in part by protein affinity columns, coimmunoprecipitation experiments, and cross-linking, since all the proteins in the applied extract have equal opportunity to bind. It is not addressed in affinity blotting or library-based detection methods, in which gene products are tested individually.

Influence of cofactors. Two types of cofactors can influence protein-protein interactions. First, small effector molecules and ions such as ATP, GTP, and Ca^{2+} can influence many protein-protein interactions. Second, other macromolecules (DNA, RNA, and proteins) can affect protein-protein interactions by forming ternary (or larger) complexes. Such complexes can be very much more stable than the corresponding binary complexes.

Effect of cellular compartmentation. A protein that is interacting with a ligand or a set of ligands is also influenced by its location in the cell. For example, some transcription factors are regulated in part by their partitioning between the cytoplasm and nucleus; they can interact with the transcription machinery only when they are in the nucleus.

Solution conditions. Other factors that can affect the strength of protein-protein interactions include solution conditions (salt concentration, pH, etc.), as well as the effects of molecules such as polyethylene glycol, which causes macromolecular crowding and can significantly lower the observed binding constant of proteins (see, for example, reference 108).

Limits of Binding-Constant Considerations

The lower limit for the concentration of a protein in an organism of the size of the yeast *S. cerevisiae* is 0.1 nM (as-

TABLE 1. Dissociation constants for some well-defined protein-protein interactions

Complex[a]	K_d (M)	Method[b]	Reference(s)
PDEαβ:PDEγ	1.3×10^{-10}	Activity	103
	5×10^{-11}	fl. an.	25
	1×10^{-11}	fl. an.	234
	$<1 \times 10^{-11}$	Activity	233
TαGTPγS:PDEγ	$<1 \times 10^{-10}$	int. fl.	164
TαGDP:PDEγ	3×10^{-9}	int. fl.	164
CAP cAMP:RNA polh	3×10^{-5}	fl. an.	91
	1×10^{-6}	fl. an.	170
T7 gene 2.5 protein:T7 DNA polymerase	1.1×10^{-6}	fl. an.	115
λ repressor (dimer to tetramer)	2.3×10^{-6}	fl. an.	9
λ repressor (monomer:dimer)	2×10^{-8}	l.z. gf.	13, 193
Citrate synthase: malate dehydrogenase	1×10^{-6}	fl. an.	214
C4 binding protein: human protein S	6×10^{-10}	Solid phase	158
p85 (PI3K): tyrosine-phosphorylated peptide from PDGF	5.2×10^{-8}	SPR	166
CheY:CheA	3×10^{-8}	SPR	197
CheA:CheW	1.3×10^{-5}	eq. gf.	72
VAMP2:syntaxin A	4.7×10^{-6}	SPR	27
EGF:EGF receptor	4.1×10^{-7}	SPR	249
PKA-C:PKA-R	2.3×10^{-10}	SPR	88
PRI:angiogenin	7×10^{-16}	Fluorescence. exch	126
ras:raf	5×10^{-8}	GST ppt'n	227
NusB:S10	1×10^{-7}	Sucrose gradient sed'n	138
NusA: core RNA polymerase	1×10^{-7}	Sucrose gradient sed'n	80
		Fluorescence tag	76
Trypsin:pancreatic trypsin inhibitor	6×10^{-14}	Kinetics, comp'n	221

[a] Abbreviations: PDE, phosphodiesterase; TαGTPγS, α subunit of transducin complexed with GTPγS; TαGDP, α subunit of transducin complexed with GDP; CAP cAMP, catabolite gene activator protein complexed with cAMP; RNA polh, RNA polymerase holoenzyme; PDGF, platelet-derived growth factor; VAMP2, vesicle-associated membrane protein 2; PKA-C, catalytic subunit of protein kinase A; PKA-R, regulatory subunit of protein kinase A.

[b] Abbreviations: fl. an., fluorescence anisotropy; int. fl., intrinsic fluorescence; l.z. gf, large zone equilibrium gel filtration; eq. gf., equilibrium gel filtration; SPR, surface plasmon resonance; exch, exchange; ppt'n, precipitation; sed'n, sedimentation; comp'n, competition.

suming a radius of 1.5 μm and one molecule per cell), and for an animal cell with a radius of 10 μm, the lower limit is about 0.3 pM. Thus, for two such proteins to interact a significant percentage of the time, the dissociation constant must be at the same concentration (in which case they will interact 38% of the time). At the other extreme, some glycolytic proteins represent 1% or more of the soluble protein in the cell. Indeed, glyceraldehyde-3-phosphate dehydrogenase is reported to approach 20% of the soluble protein in *S. cerevisiae* under certain conditions. This upper limit corresponds to 1.7×10^7 protein molecules per cell and a cellular concentration of 1 mM, and it represents the upper limit for binding-constant considerations of two such proteins. In considering protein concentrations, it is worth noting that a typical yeast cell contains about 3×10^5 ribosomes (226), 100 to 500 molecules of tRNA splicing enzymes (169, 178), and 300,000 molecules of actin (157).

Methods for Determining Binding Constants

A number of methods have been described to measure binding constants. Some of the more commonly used ones are described below, together with a brief evaluation of the method. The values of dissociation constants for several protein-protein interactions are listed in Table 1.

Binding to immobilized proteins. Protein affinity chromatography can be used to estimate the binding constant. This method is well described in an excellent review (69). The form of the binding equation that is used in this sort of experiment expresses the fraction of L bound to protein P as follows: $[PL]/[L_t] = [P_f]/([P_f] + K_d)$. As long as the concentration of covalently bound protein $[P_t]$ is in great excess over that of the ligand, $[P_t] \approx [P_f]$ and the fraction of protein L that is bound is $[P_t]/(K_d + [P_t])$. Thus, if $[P_t] = 100\ K_d$, essentially all of L is bound (a little more than 99%), and if $[P_t] = 0.01\ K_d$, very little of L is bound (a little less than 1%).

Columns are prepared with different concentrations of covalently bound protein. Then a preparation of the interacting protein ligand is loaded on the column and washed with 10 column volumes of buffer, and bound protein is eluted with SDS. At a concentration of 20 K_d, the covalently bound protein retains 95% of the ligand in one column volume and therefore 0.95^{10} or 61% in 10 column volumes. Thus, the lowest concentration of bound protein that allows retention of most of the ligand is 20 K_d.

The percentage of bound ligand drops very quickly as the concentration of covalently bound P on the column is lowered, particularly as the concentration of P_t approaches K_d. At 5 K_d 16% of the ligand would be retained, at 2 K_d 1.7% of the protein would be retained, and at 1 K_d only 0.1% would be retained. It is for this reason that detection of interacting proteins by affinity chromatography depends critically on the concentration rather than the amount of bound protein (see the section on protein affinity chromatography, above).

An important parameter in this experiment is the amount of protein that is active on the column. Estimates range from 10% for gene 32 protein to about 50% for others (69). A second factor is the amount of pure protein available to be coupled. If protein is limiting, sufficiently high concentrations of bound protein on the gel are achieved only with appropriate microcolumns. Such columns, with as little as 20 μl of appropriate beads, are described in detail by Formosa et al. (69). With the recent widespread use of gen fusion technology, large quantities of protein are not a serious problem with most cloned structural genes. A third factor, which is evident from the discussion above, is the form of the protein that is used for the determination. Proteins that require modification to be active must be purified in that form for proper evaluation.

This method works well in estimating the binding constant. However, it is not clear that the values obtained represent a true equilibrium constant; if so, one would have to assume that the bound ligand is always in equilibrium with the solution ligand during flow of the column and that interactions of solid-phase bound protein with liquid-phase ligand are the same as interactions in the liquid state. Nonetheless, for interactions that have been measured by more than one method, the results agree well (see reference 69 and references therein).

Sedimentation through gradients. The method of sedimentation through gradients measures populations of complexes by monitoring the rate of sedimentation of a mixture of proteins through gradients of glycerol or sucrose. Fractions are assayed by appropriate methods (activity, immunoblotting, etc.) to determine the elution positions of each protein. Proteins will sediment as a complex at concentrations above the binding constant (provided that the complex is stable; see the discussion below) and at their native positions at concentrations below the binding constant. By varying the concentration of one or both of the proteins and taking into account the dilution of the species during sedimentation, one can reasonably accurately bracket the binding constant. For example, the binding constant of *E. coli* NusB protein and ribosomal protein S10 was estimated at 10^{-7} M based on the observation that S10 protein sedimented faster (with NusB protein) when both were at 6×10^{-7} M, slightly more slowly when both were at 3×10^{-7} M, and much more slowly (midway between its sedimentation position alone and its fully complexed sedimentation position) when both were at 1.5×10^{-7} M (138). There are two reasons that S10 sedimented at an intermediate position rather than at its own position during the run at 1.5×10^{-7} M of each protein. First, the proteins are usually about fivefold more dilute at the end of the sedimentation than when they are first loaded on the gradient; therefore, if S10 protein could bind at the beginning of the run (and sediment faster), it might not bind at the more dilute concentration at the end of the run. Thus, it would sediment at an intermediate position. Second, equilibrium binding is a dynamic process and molecules are constantly associating and dissociating. Therefore, an individual S10 molecule which dissociated from NusB at the trailing edge of the peak would be in a region with very much less NusB to bind. It would sediment at its native rate from that point on.

There are two problems associated with this technique. First, it is not an equilibrium determination, because of the changing conditions during the run. Therefore, failure to detect an interaction may be due to rapid equilibrium rather than a lack of interaction. As such, values obtained from this type of experiment represent an upper bound for the binding constant. Second, sedimentation through gradients does not resolve species that well. Sedimentation rates vary as $M^{2/3}$ for spherical molecules. Thus, dimerization of one spherical molecule with one that is 1/10 the mass will increase its sedimentation rate by only 6%, which is very difficult to detect; in contrast, the change in mobility of the smaller molecule will be fivefold under such conditions.

Although this method has limitations, it has been useful for estimating the upper limit of a binding interaction.

Gel filtration columns. Gel filtration is another simple way of estimating the binding constant. In gel filtration, the elution position of a protein or of a protein complex depends on its Stokes radius. This provides a very powerful and conceptually simple method for evaluating the strength of the interaction between two different proteins. Such sizing columns have been used in three distinct ways to measure or to estimate the binding constant.

(i) Nonequilibrium "small-zone" gel filtration columns. In the simplest approach, a solution containing a protein and a ligand protein is applied in a small volume to the column and the material is resolved in the usual way. This is called a "small-zone" column. The elution positions of the protein and ligand in the mixture are compared with those of the protein and ligand when each is chromatographed individually on the same column. If a complex has formed between the protein and ligand, the complex will elute earlier than either protein alone. From measurements of the concentrations of species required to form a complex, one can estimate the binding constant. This type of experiment has been used, for example, to measure the binding of *E. coli* NusA protein to core RNA polymerase and has yielded values very similar to those determined by fluorescence measurements (76). Similarly, Herberg and Taylor (89) quantitated the interaction of cAMP-dependent protein kinase with both the R1 subunit and PKI in the presence and absence of MgATP.

This direct-application method is not an equilibrium method. Since the concentrations of species change during gel filtration (by diffusion and by dilution), the results are subject to the same sources of error as those of sedimentation through sucrose gradients (see references 2 and 250 for a discussion). Thus, the binding constants calculated in this way can be vastly underestimated, particularly if the complex is in rapid equilibrium (see Fig. 3 of Gegner and Dahlquist [72]) for a vivid contrast between nonequilibrium and equilibrium gel filtration). However, several modeling systems have been described (see reference 211 and references therein).

(ii) Hummel-Dreyer method of equilibrium gel filtration. Gel filtration can also be used as an equilibrium method to establish the binding constant between a protein and its ligand protein. One such method is based on the classic paper by Hummel and Dreyer (102). In this gel filtration method, both the gel filtration buffer and the sample had ligand at the same concentration, but only the sample contained protein. Elution of a protein through such a column caused an increase in the concentration of ligand where the protein eluted, followed by a trough of ligand concentration representing ligand that had been removed in the binding. Evaluation of the binding constant of the protein-ligand complex was simply a matter of knowing the concentration of protein eluted, the free concentration of ligand (set by the column), and the concentration of ligand bound with protein (the concentration of ligand in samples containing protein).

This elegant method has been applied to the interaction of two proteins in only a few cases. As illustrated in Fig. 10, the gel filtration buffer contains protein ligand, and the applied sample contains gel filtration buffer (with the same concentration of protein ligand) as well as the other protein. Gegner and Dahlquist (72) used a column equilibrated with CheW to demonstrate and quantitate the interaction of CheA with CheW. They varied the CheW concentration in the initial sample (while maintaining a constant concentration of CheA in the sample and CheW in the buffer) and quantitated the peak area at the CheW position. The CheW concentration in the sample at which there was no resulting CheW peak or trough represented a sample at true equilibrium. From this, they could calculate a dissociation constant of the complex of 13 μM. A similar series of experiments was done by Yong et al. (243) to demonstrate an interaction between glycerol-3-phosphate dehydrogenase and lactate dehydrogenase over an extremely limited range of NADH concentrations. Such a complex was observed only when the NADH concentration was high enough for an interaction and low enough to be shared by the two enzymes, and it provided evidence for substrate channeling.

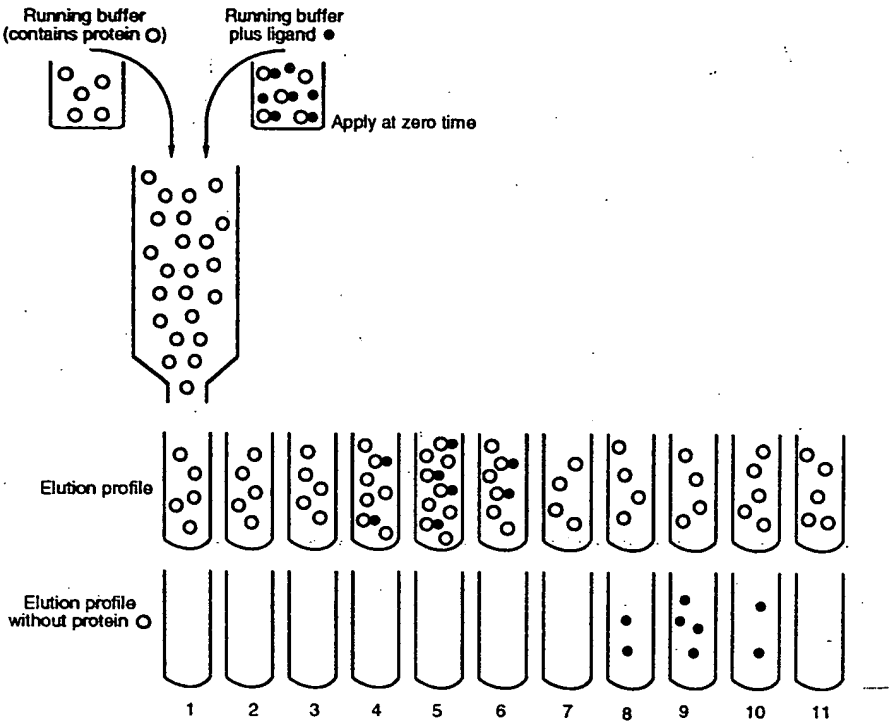


FIG. 10. Equilibrium gel filtration. A solution containing both protein ligand (solid circles) and interacting protein (open circles) is applied to a gel filtration column which is equilibrated with solution containing the interacting protein and developed with running buffer containing the interacting protein. The elution pattern is shown in the first row of test tubes at the bottom. The second row of test tubes indicates the elution pattern that would be observed in the absence of interacting protein.

This method is so simple and inexpensive that it is likely to become much more widely used than at present. Moreover, as an equilibrium experiment, it is without any flaw. The only requirements of this technique are that sufficient protein is available for the experiments and that the elution position of the complex differs from that of at least one of the interacting proteins. With the development of rapid techniques for large-scale protein purification through the use of fusion proteins, it should become relatively routine to obtain enough of any protein to use as a column eluant.

Another variation of Hummel-Dreyer columns is the partitioning method. In this technique, a protein and its ligand protein are mixed with a gel and allowed to equilibrate and the gel is centrifuged or filtered to separate the aqueous phase. From an analysis of the distribution of the protein and the ligand protein in the filtrate and in the gel when they are added separately or together, the K_d can be calculated. An example of this technique is the demonstration of a complex between transaminase and glutamate dehydrogenase which occurs with a dissociation constant of 16 to 67 μM, depending on the presence of various metabolites (63); this is another example of metabolite channeling. This method is also not in wide use, although it seems simple and accurate.

(iii) Large-zone equilibrium gel filtration. One final method of equilibrium gel filtration is the large-zone method (1, 2), in which a very large sample volume is applied to the column, followed by conventional buffer elution. Because a large volume is applied, the concentration of the eluted protein is fixed and constant during the experiment, except at the leading and trailing edges. The elution position of the leading or trailing edge (which measures the size of the molecule) is then monitored as a function of the sample concentration applied to the column. From such experiments, calculation of the dissociation constant is thermodynamically rigorous, as it is for the Hummel-Dreyer method. This large-zone method has been used to monitor self-association of proteins as well as interactions of dissimilar subunits (see, for example, references 75 and 122), but it has received only limited attention because of the large amounts of protein needed to do the experiments.

A variation of this method, first described by Sauer (193), monitors the change in elution position of radiolabeled protein mixed with different concentrations of unlabeled protein in different runs. The use of labeled protein allows simpler and more accurate determination of the elution position, thus allowing Sauer to determine a dimerization constant of 20 nM for repressor. Improvements in protein labeling have demonstrated that the lower limit of detection for this method is a K_d of the order of 10^{-12} M (13).

Sedimentation equilibrium. Although sedimentation equilibrium is a classical method of determining the molecular weight of a protein, it has not been widely used to study protein-protein interactions. However, recent progress makes this method much more accessible on a day-to-day basis (see reference 185 for a recent review). Sedimentation equilibrium can now be done in everyday preparative ultracentrifuges with swinging-bucket rotors, and samples can be readily collected because of the development of a highly reproducible BRANDEL microfraction collector (183). These developments allow the use of a variety of techniques to assay the protein content of each sample, including kinetic assays, radioactive tracers (183), and gel analysis of samples (47); the result is a huge increase in sensitivity over that obtained with the old model E centrifuge (184).

Fluorescence methods. Since fluorescence is a highly sensitive method for detecting proteins through their tryptophan residues, it is potentially a useful way of evaluating protein-protein interactions. Two such methods have been used and are described below.

(i) Fluorescence spectrum. Changes in the fluorescence emission spectrum on complex formation can occur either by a shift in the wavelength of maximum fluorescence emission or by a shift in fluorescence intensity caused by the mixing of two proteins. Therefore, the fluorescence intensity at a particular wavelength can be used to evaluate the dissociation constant. A good example of this technique is illustrated by the interaction of the γ subunit of cGMP phosphodiesterase (PDEγ) subunit with the transducin α subunit (Tα) in the presence of GTPγS or GDP (164).

An equimolar solution of TαGTPγS and PDEγ causes a blue shift in the fluorescence emission spectrum relative to the sum of the individual fluorescence spectra, resulting in a difference spectrum [F (complex) − F (sum)] with a positive component at low wavelengths (320 nm) and a negative component at higher wavelengths (357 nm). Titration of PDEγ into a solution of TαGTPγS therefore caused an enhanced increase in the fluorescence at 320 nm relative to that observed by titration of PDEγ into buffer alone (and a corresponding decrease at 357 nm) until the TαGTPγS was all complexed, after which further addition PDEγ caused no changes in fluorescence intensity relative to that observed in buffer alone. When corrected for PDEγ fluorescence, both curves yielded the same binding curve, and the K_d for the interaction was evaluated at <100 pM. The interaction of TαGDP with PDEγ results in a large increase (ca. 70%) in the intensity of the fluorescence emission spectrum relative to the sum of the individual spectra, and this was used to evaluate the K_d at 2.75 nM.

This technique has two limitations. First, the probability of detecting a change in the fluorescence spectrum decreases with the total number of tryptophan residues in the two proteins, since the fluorescence spectrum is the sum of the contributions from all the tryptophan residues. Since PDEγ has only one tryptophan residue and Tα has two, this condition was easily met in studying the Tα-PDEγ complex. Second, the sensitivity is limited by the intensity of the fluorescence change, which in turn depends on the inherent sensitivity of fluorescence (of the order of nanomolar) and the change that is observed (which is not easily predictable). Thus, the binding constant was too low to evaluate the TαGTPγS-PDEγ interaction (<100 pM) but was high enough to evaluate the interaction in the presence of GDP (2.75 nM).

Although these two limitations exclude the study of many interactions, a number of proteins have a small or limited number of tryptophan residues. For example, bovine Hsc70 has only two tryptophans, and its interaction with small peptides has been evaluated because of the resulting quenching of the fluorescence intensity (123). Similarly, the interaction of angiogenin (one tryptophan) with human placental RNase inhibitor (six tryptophan residues) causes a 50% increase in fluorescence (126), and the dissociation of mitochondrial creatine kinase (four tryptophans per monomer) from octamers to dimers results in a 25% decrease in fluorescence (81).

A second way in which fluorescence is used to measure the interaction of proteins is with a fluorescent tag. This allows for greater sensitivity of monitoring interactions, as long as the fluorescent adducts do not adversely affect the function of the modified protein or its interaction with other proteins. An example of this approach is the interaction of spinach calmodulin with smooth myosin light-chain kinase (146). Calmodulin from spinach has a single cysteine, which could be quantitatively labeled with 2-(4-maleimidoanilino)-naphthalene-6-sulfonic acid (MIANS). Calmodulin labeled with MIANS was as efficient as the wild type in activating calcineurin, in activating cGMP-dependent phosphodiesterase, and in binding terbium. The fluorescence of MIANS-labeled calmodulin increased 80% on binding calcineurin, more than fourfold when bound with myosin light-chain kinase, and twofold on binding caldesmon. In each case, the fluorescence change required the presence of calcium, and titrations were done to measure the K_d (<5, 9, and 250 nM, respectively).

(ii) Fluorescence polarization or anisotropy with tagged molecules. Because of the long lifetimes of excited fluorescent molecules (nanoseconds), fluorescence can also be used to monitor the rotational motion of molecules, which occurs on this timescale. This is accomplished experimentally by the use of plane-polarized light for excitation, followed by measurement of the emission at parallel and perpendicular planes. Since rotational correlation times depend on the size of the molecule (approximately 1 ns/2,400 Da for an idealized molecule), this method can be used to measure the affinity of two proteins for one another because of the increased rotational correlation time of the complex. Fluorescence anisotropy is done most often with a protein bearing a covalently added fluorescent group, which increases both the observed fluorescence lifetime of the excited state and the intensity of the fluorescent signal.

A good example of this technique is described by Weiel and Hershey (229), who studied the interaction of protein synthesis initiation factor 3 (IF3) with 30S ribosomal subunits by using fluorescein-labeled IF3. The labeled protein routinely had about one dye molecule per monomer, and most of the IF3 protein had one or two dye molecules attached. Fluorescein-labeled IF3 was biologically functional: it bound 30S ribosomal subunits, as measured by sucrose density gradients, at a saturable site(s) and had 80 to 100% of the activity of the native protein in stimulating binding of $tRNA^{Met}$ to 70S ribosomes in the presence of RNA. In the presence of 30S ribosomes, both the fluorescence emission spectrum and the fluorescence lifetime of the fluorescein-labeled IF3 were unchanged. Thus, the observed increase in fluorescence polarization which was associated with binding of 30S ribosomes was most consistent with the expected change in polarization as a result of binding a larger molecule. The Scatchard plot derived from the polarization data gave a stoichiometry of 1:1, and the dissociation constant from the polarization data was 3.2×10^{-8} M. Moreover, wild-type nonderivatized IF3 competed for the binding site with the same binding constant. Thus, the fluorescent probe had no effect on any measurable parameter and the measured K_d is likely to be accurate.

Similar experiments have been done with a variety of systems to evaluate the strength of protein-protein interactions. Fluorescein-labeled IF2 was slightly less active than nonderivatized protein, and the binding to 30S ribosomes was twofold weaker than that of the corresponding unlabeled protein (230). T7 gene 2.5 protein labeled with near-molar amounts of fluorescein isothiocyanate caused both a decrease in fluorescence and an increase in anisotropy when bound with T7 DNA polymerase. The fluorescein isothiocyanate-modified protein had no effect on activity, and the binding constant determined by anisotropy (1 μM) was nearly the same as that determined by anisotropy measurements of EDANS-labeled gene 2.5 protein (1.3 μM), for which the rotational correlation time indicated a 1:1 complex (115). The interaction of (florescein-labeled) citrate synthase and malate dehydrogenase was shown to be well within the physiological range (K_d = 1 μM) and varied as much as 25-fold in the presence of different metabolites (214). The tetramer-dimer equilibrium of λ repressor could be observed with dansylated λ repressor, because of its long fluorescence lifetime and high anisotropic value (indicating rigid orientation), but not with fluorescein, which was attached in the highly

mobile N-terminal arm of the repressor molecule (and therefore gave low values) (9).

A variation of this technique has been developed for the interaction of a DNA-binding protein with another protein, in which the DNA is fluorescently labeled (91). In this way, *E. coli* CAP could be shown to interact with RNA polymerase holoenzyme in the presence of cAMP and in the absence of a promoter site. The fluorescently labeled DNA oligonucleotide had a CAP-binding site but no RNA polymerase-binding site, and the resulting increase in polarization allowed the determination of a CAP-RNA polymerase binding constant (2.8×10^{-7} M). Since this interaction was not observed with a CAP mutant protein that was defective in transcription activation, it seems likely that the interaction is important physiologically. Other fluorescent polarization experiments suggest that the CAP-RNA polymerase interaction is much stronger in the presence of cAMP and requires σ factor (170).

Solution equilibrium measured with immobilized binding protein. A simple technique for measuring the dissociation constant of a solution of interacting proteins makes use of bound competitor protein to determine the amount of free protein in such a solution. This method was first described for antibody-antigen reactions (71) and later modified for general use to determine the interaction of C4b-binding protein (C4BP) with human protein S (HPS) (158). A solution containing C4BP and HPS was incubated until equilibrium was reached. The amount of free C4BP in the solution was then determined by incubating an aliquot on a plate containing immobilized HPS under conditions (short incubation time) in which a limited amount of the free C4BP binds the immobilized HPS. This resulted in little perturbation of the equilibrium during the assay for C4BP retained by the immobilized HPS, which was quantitated by an antibody-based method.

This method requires satisfaction of three criteria. First, the two proteins (HPS in solution and HPS immobilized on the plate) cannot bind each other. If they did, C4BP could be captured through HPS-HPS interactions. Second, HPS in solution and HPS immobilized on the plate must compete for the same binding site. This is obviously true in this case, but it is not necessarily true if, for example, anti-C4BP is used in the immobilized system to detect the amount of free C4BP. Third, the method requires that only free C4BP be measured during the incubation with immobilized HPS. This in turn requires that binding to the immobilized HPS remove only a small portion of the total C4BP (<10% was removed in this example) so that equilibrium of the solution is perturbed as little as possible. This condition also requires that the off rate of the complex is low compared with the time of incubation with the immobilized HPS; otherwise, HPS-C4BP complexes could dissociate during the incubation with immobilized HPS and the dissociated C4BP would be measured as free C4BP. Thus, this method, although simple, provides only an upper bound of the dissociation constant.

Surface plasmon resonance. The recent development of a machine to monitor protein-protein and ligand-receptor interactions by using changes in surface plasmon resonance measured in real time spells the beginning of a minor revolution in biology. This method measures complex formation by monitoring changes in the resonance angle of light impinging on a gold surface as a result of changes in the refractive index of the surface up to 300 nm away. A ligand of interest (peptide or protein in this case) is immobilized on a dextran polymer, and a solution of interacting protein is flowed through a cell, one wall of which is composed of this polymer. Protein that interacts with the immobilized ligand is retained on th polymer surface, which alters the resonance angle of impinging light as a result of the change in refractive index brought about by increased amounts of protein near the polymer. Since all proteins have the same refractive index and since there is a linear correlation between resonance angle shift and protein concentration near the surface, this allows one to measure changes in protein concentration at the surface due to protein-protein or protein-peptide binding. Furthermore, this can be done in real time, allowing direct measurement of both the on rate and the off rate of complex formation. A good layman's review of surface plasmon resonance is found in articles by Malmqvist (136) and Jonsson et al. (109), and a clear derivation of the appropriate equations is found in the article by Karlsson et al. (111).

In practice, determination of a binding constant requires measurement of two parameters. First, the increase in RU (resonance units) is measured as a function of time by passing a solution of interacting protein past the immobilized ligand until (usually) the RU values stabilize. Second, the decrease in RU is measured as a function of time with buffer lacking interacting protein. This produces a sensorgram for each concentration of protein, a continuous recording of RU versus time. This procedure is then repeated at a number of protein concentrations, after regeneration of the dextran surface. From these two sets of data, two lines are constructed whose slopes correspond to k_a (the on rate) and k_d (the off rate); from these data, K_d is calculated as k_d/k_a. An alternative determination of K_d can be made by using the steady-state RU values at different protein concentrations.

This system has several advantages. First, it requires very little material. Typically only 1 to 10 μg of protein has to be immobilized on a sensor chip, which can be reused up to 50 times after removal of adhering protein. Similarly, solutions of interacting protein are in the range of 0.01 to 1 ml, depending on the chosen flow rate (109). Second, the method is very fast. A typical run for a given protein takes about 10 min. Third, no modifications of the proteins are required, such as labeling or fluorescent tags. Fourth, interactions can be observed even in complex mixtures. Fifth, both the on rate and the off rate are readily obtained. Sixth, the system is useful over a wide range of protein concentrations. The practical lower limit of the original Biacore system is a change in resonance angle of 10^{-3} degrees (10 RU), corresponding to surface concentrations of 10 pg/mm^2; moreover, the system is linear up to RU values of 30,000 (109). Seventh, the system is quite sensitive; the practical limit for association rates is 10^6/M/s, and off rates as low as 1.1×10^{-5}/s have been measured by recording for 6 h with buffer (197).

This technique has been used successfully to monitor protein-peptide interactions. A good example is the determination of the binding interaction of different SH2 domains with two tyrosine-phosphorylated substrate peptides derived from platelet-derived growth factor (166). The corresponding peptides were attached to the dextran polymer chip via avidin on the chip and biotin on the peptides. Subsequent real-time analysis demonstrated that interaction of these peptides with the p85 subunit of phosphatidylinositol-3-kinase (PI3K) was characterized by a very high association rate (2×10^6/M/s) and dissociation rate (0.1/s) for the 12-mer peptide Y740P and that most of this binding was contributed by the C-terminal subunit of p85. In this particular case, the dissociation rate of bound p85 had to be determined in the presence of a sink of excess competing peptide in the buffer; otherwise, rebinding of dissociated p85 was a significant problem because of the very high on rate. A similar study of p85 SH2 domain interactions with different tyrosine-phosphorylated peptides (from IRS-1) led to the same conclusions of a high on rat and off rat , which was

also measured in the presence of a sink of peptide (64). In this case, the on rate was too high to measure directly (as high as 4.4×10^8/M/s for the C-terminal SH2 domain of p85) and was instead inferred from steady-state binding and off rate measurements and confirmed by competition experiments with free phosphorylated peptide (64). On rates in excess of 10^6/M/s can be limited by mass transport rates (fluid flow through the cell) rather than binding-reaction rates, although this can be partially compensated by either higher flow rates or a smaller amount of peptide on the chip (111). Competition experiments were also used to show that the affinity of p85 for phosphorylated peptides was 300- to 800-fold greater than for the corresponding nonphosphorylated peptide and was as much as 100-fold weaker with a glycine or arginine at the +1 position relative to the tyrosine compared with bulky hydrophobic groups or glutamate (64).

One final study demonstrated that a specific threonine residue in the SH2 domain of Src, when changed to a tryptophan, increased the affinity of the domain for phosphorylated peptides which were substrates for GRB2 and that the corresponding tryptophan of GRB2, when altered to threonine, weakened the affinity of GRB2 for this peptide (137). In each of these three examples, the primary determinant of specificity was the on rate rather than the off rate.

Surface plasmon resonance has also been used with great success to monitor protein-protein interactions. One such example is the demonstration of a quarternary complex of CheY with CheA, CheW, and Tar (197). CheY was bound to the dextran surface through a unique (and engineered) cysteine residue, which did not affect chemotaxis activity and which was remote from the interaction domain (197). CheA binds this immobilized CheY protein with a low association rate (368/M/s) and a very low off rate (1.14×10^{-5}/s). Moreover, CheA, CheW, and Tar probably form a quaternary complex with CheY; addition of all three proteins greatly increases the amount of protein bound to CheY relative to that obtained with CheA alone, although neither Tar nor CheW binds CheY individually or when present together.

Other examples of protein-protein interactions studied by surface plasmon resonance include the interaction of monoclonal antibodies with human immunodeficiency virus type 1 core protein p24 (111), EGF with the EGF receptor (249), the regulatory and catalytic domains of cAMP-dependent protein kinase (88), and VAMP2 and sytaxin 1A (27).

Two minor problems are associated with surface plasmon resonance measurements. First, immobilization of the ligand protein must be of such a nature that it does not impede or artificially enhance interactions. This is the same problem that is associated with protein affinity columns. Attachment of CheY was accomplished by using a single site remote from the interaction domain (197); this presents the interacting face to the solvent. Phosphorylated peptides were attached by biotinylation of the peptide at a single site (but variable position) with a long spacer followed by noncovalent interaction with an avidin-coupled sensor chip (166), and attachment of monoclonal antibodies to the chip was accomplished through noncovalent binding to covalently coupled rabbit anti-mouse IGGFc (111). Primary amines are often linked directly to the dextran polymer, leading to more homogeneous presentation of surfaces to the solvent but causing possible inhomogeneities in interaction (88). Second, the sensor chip has to be regenerated under conditions which do not denature the immobilized ligand protein. Protein adhering to the immobilized C subunit of prot in kinase A was removed with cAMP (88), proteins binding to immobilized phosphorylated peptides were removed with a pulse of dilute SDS (166), and CheY was regenerated with a pulse of guanidine hydrochloride (197). In some cases, the ligand is deliberately removed before the next experiment; thus, monoclonal antibodies sticking to IGGFc were removed with dilute HCl before readdition of the monoclonal antibodies to act as a ligand for p24 binding (111).

Limits to Detection

Determination of the binding constant of tightly interacting species by standard methods described above depends on being able to determine and quantitate the fraction of protein ligand bound at a given protein concentration that spans the dissociation constant. For a standard 50,000-kDa protein, the practical limit of silver staining is of the order of 0.2 ng or 20 μl of a 10-ng/ml solution, which would be useful for a dissociation constant of 1 nM or greater. For in vitro translated protein, the practical limit is 1,000 Ci/mmol times the number of amino acid residues, or 1,000 dpm of ^{35}S-labeled protein per fmol (singly labeled); this corresponds to 10^{-12} M or, with 10 residues incorporated, 10^{-13} M; therefore, allowing for concentrations below K_d, the lower limit of detection is of the order of 10^{-12} M.

Some protein-protein interactions are too tight ($K_d < 10^{-12}$ M) to measure by the methods described above. For example, human placental RNase inhibitor (PRI) interacts very tightly with both angiogenin ($K_d = 7 \times 10^{-16}$ M) (126, 126a) and human placental RNase ($K_d = 9 \times 10^{-16}$ M) (199). For the interaction of PRI with angiogenin, the association rate constant, k_a, was measured by monitoring the change in intrinsic fluorescence by stopped-flow fluorescence techniques, and the dissociation rate constant, k_d, was measured by measuring the release of PRI in the presence of scavenger RNase, to which it binds and inhibits the activity.

A dissociation constant of the magnitude of 7×10^{-16} M for the PRI-angiogenin interactions means that the dissociation rate is measured in weeks! In this case, the $t_{1/2}$ for dissociation of the complex was 60 days (corresponding to $k_d = 1.3 \times 10^{-7}$/s). Furthermore, the overall on rate of 1.8×10^8/M/s liters · mol/s is near the diffusion limit for molecules of the size of proteins. It is hard to imagine what selective pressure would require or maintain such a tight interaction. This is particularly true since human placental RNase and angiogenin both bind PRI equally tightly and are substantially different at the amino acid level.

It is possible that a number of macroscopic protein-protein interactions operate at this level. Any protein composed of three or more subunits can have significant interactions among individual pairs of the component protein. If, for example, a subunit has a K_d of 10^{-7} M with each of two other subunits, the effective K_d of the dissociation of that subunit from the complex is 10^{-14} M (see reference 116 for a discussion of this point). Thus, complicated structures like the ribosome might effectively lock the proteins together in undissociable units. It is also possible that other, simpler interactions are this tight; the dissociation rate of the subunits of a number of proteins that purify as a complex tends never to be investigated.

EXAMPLES OF WELL-CHARACTERIZED DOMAINS

Given that a straightforward set of experiments is all that is required nowadays to identify two proteins that interact and to delineate the domains responsible for the binding, toward what ends does this analysis continue? To address this question, it is instructive to consider the case of some domains involved in protein-protein interaction that have been extensively characterized. Using a combination of numerous techniques, includ-

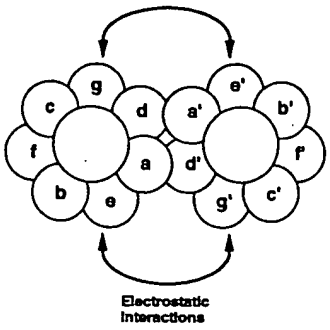


FIG. 11. Helical wheel representation of a leucine zipper. Adapted from reference 221a with permission of the publisher.

ing detailed structural approaches, investigators who have focused on the analysis of leucine zippers, SH2 domains, and SH3 domains have made tremendous advances in the last few years. These studies have considerably extended our understanding of transcriptional regulation and signal transduction. In the next sections, we provide a brief view of how these three domains function.

Leucine Zipper

The leucine zipper is a protein-protein interaction motif in which there is a cyclical occurrence of leucine residues every seventh residue over short stretches of a protein in an α-helix. These leucine residues project into an adjacent leucine zipper repeat by interdigitating into the adjacent helix, forming a stable coiled-coil. This motif was first described by Landschulz et al. (124) in connection with a new structure within DNA-binding proteins that might be responsible for interactions with a similar motif to promote specific DNA binding by basic amino acid residues adjacent to the leucine zipper motif (hence the name bZIP). The leucine zipper model was originally proposed on the basis of the leucine distribution and amino acid sequence of regions of C/EBP, Myc, Fos, Jun, and Gcn4. It is now known to be common to over 30 proteins (59). Subsequent experiments have confirmed the existence of this structure and have extended these observations.

Structure. The X-ray structure of the Gcn4 leucine zipper region (consisting of 33 amino acids) demonstrates that the leucine zipper consists of two parallel coiled coils of α-helices wrapped around each other and forming one-quarter of a turn of a left-handed supercoil (59, 161; also see reference 4). The dimer forms a smoothly bent cylinder about 45 Å (4.5 nm) long and 30 Å (3 nm) wide. On a helical-wheel representation of the α-helix (Fig. 11), the leucines occupy position d (and d′ of the adjacent helix) and share the interior with the residues at position a (a′), as well as parts of residues e and g (and e′ and g′). The packing corresponds to the "knobs into holes" model proposed by Crick (42), in which each interior amino acid residue is packed into a gap formed by four nearest neighbors from the opposite helix. More than 95% of the surface area that is buried upon dimerization is from the side chains of these residues.

Stability. The leucine zipper coiled coil is stabilized because of three factors: the hydrophobic groups that are buried (leucines at position d and hydrophobic or neutral residues at position a); constancy of size of the internally packing residues at each position; and several distinct ion pairs. Three such ion pairs appear to form, and each is between the e of one heptad and the g of the other. The leucine residues are critical for function in Gcn4. Although each individual leucine can itself be replaced by several different hydrophobic residues, randomized substitution of the leucines with other hydrophobic residues invariably causes the protein to lose function when more than one leucine is substituted; furthermore, isoleucine is by far the most easily tolerated substitution (98).

The binding constant of leucine zipper moieties that interact is estimated to be in the nanomolar range (163) and has been measured at 5×10^{-8} M for the Jun-Jun dimer at 4°C (196). Even a peptide corresponding to the Fos leucine zipper, which does not dimerize in vitro, has been shown to dimerize in the micromolar range (163).

The leucine zipper moieties that naturally interact do not necessarily have the maximal stability. For example, the Gcn4 dimer has a buried asparagine residue which is present within the hydrophobic core (59, 161). This Asn residue packs loosely in the crystal structure, and this position is particularly tolerant of other amino acids (98). Moreover, the asparagine residue (and resultant internal hydrogen bond) drastically destabilizes the Gcn4 zipper; its replacement with valine stabilizes the coiled coil about 1,000-fold (28). It has been speculated that the internal asparagine of Gcn4 (and, by extension, other buried polar groups in the a position in other leucine zippers) is present, so that the proteins do not bind too tightly and therefore can be subject to regulation, or that it keeps the coiled coils in register (4).

Specificity. The specificity of leucine zippers is the key to their regulatory properties. The oncoproteins Fos and Jun, for example, associate with each other to form a heterodimer in preference to the Jun-Jun homodimer. This preference has important consequences in that Fos-Jun heterodimers and Jun-Jun homodimers bend DNA in opposite orientations (114), which may explain the fact that Jun interaction with the glucocorticoid response element of the prolactin gene results in activation of the gene, whereas Fos-Jun interaction results in repression (51).

Specificity of Fos-Jun and Jun-Jun dimerization is achieved primarily by the electrostatic interactions of residues at the e and g positions at the periphery of the hydrophobic core (162). Fos has Glu residues at the g position, and Fos-Fos dimers are much more stable (as measured by T_m) at pH values at which these Glu residues are neutralized. Conversely, Jun is slightly more basic at the e and g positions, and Jun-Jun dimers are more stable at higher pH. Fos-Jun dimers, which are the preferential form, are uniformly stable over a wide range of pH values, because they are more neutral overall. A series of hybrid peptides in an otherwise Gcn4 peptide illustrate the point (162). Specificity (or antispecificity) is achieved by the 8 amino acids at the e and g positions of the peptide and not at other positions.

Regulation. Leucine zipper proteins are likely to be functionally regulated. Thus, the carboxyl-terminal zipper of the human and *Drosophila* heat shock factors may suppress formation of amino-terminal zippers in a way that is sensitive to heat shock (175). Similarly, the calphotin protein binds calcium at one end and has a distinctive leucine zipper at the other end (8). It may therefore be used to transmit signals by altering binding properties.

SH2 Domain

The SH2 domain was first recognized as a noncatalytic domain of Src that was homologous to the Fps protein (189) and is now recognized as a common motif involved in protein-protein interactions (117, 168). More than 20 SH2-containing proteins have been identified. They share a motif of about 100 amino acids that is involved in the recognition of proteins and

peptides containing phosphorylated tyrosines. This recognition is implicated in the mechanism of signal transduction, because the phosphorylated tyrosines that are recognized include those of growth factor receptors such as the platelet-derived growth factor receptor, the EGF receptor, and the fibroblast growth factor receptor. On binding their respective growth factors, the growth factor receptors have their tyrosine kinase activity activated, which allows them to autophosphorylate. The autophosphorylated receptor then binds various proteins containing SH2 domains, which are then phosphorylated to modulate their activity. Thus, the binding of growth factor on the outside of the cell results in phosphorylation on the inside of specific substrate proteins. The particular proteins that are phosphorylated depend on the binding specificity of the SH2 domains for the phosphorylated receptor. Binding of different peptides to different SH2 domains has yielded the following results.

Binding of SH2 proteins requires a large domain of the SH2 protein. The conserved domain of SH2 domains, which is common to more than two dozen proteins, has been crystallized for Src (224, 225) and solved by nuclear magnetic resonance spectroscopy techniques for c-Abl (165) and p85a of PI3K (20). In each case, this domain folds into a structure in which a set of internal antiparallel sheets is surrounded by two more or less symmetrical α-helices. The conserved amino acids tend to be part of the recognition for phosphotyrosine (e.g., Arg-175 of Src) or part of the hydrophobic pocket. Variable regions are responsible for sequence recognition (205) and may be parts of variable loops of unknown function (188).

Binding of SH2 proteins requires phosphorylated tyrosine in vitro. Thus, the binding constant of a peptide to an SH2 protein of p85 is between 50- and 800-fold weaker without the phosphate than with the phosphate (64). This preference is attributable to specific side chain contacts of the SH2 domain with the phosphoryl group of phosphotyrosine. The phosphoryl oxygens are hydrogen bonded with two guanidinium hydrogens, one from one arginine and one from another arginine, one hydroxyl hydrogen from threonine and one from serine, and a backbone amide hydrogen. One of the arginines appears to be acting both as a hydrogen bond donor and as an ion pair with the phosphate group. Thus, it cannot be substituted with lysine without loss of binding (140). These contacts are the same whether a weak-affinity (224) or a strong-affinity (225) phosphotyrosine-containing peptide is used.

SH2 domains make contacts with only a small region surrounding the phosphorylated tyrosine. Small peptides faithfully reproduce binding to SH2 domains and display binding constants of the order of nanomolar (64, 218). This is consistent with the crystallographic data of the SH2 domain of v-Src bound to a high-affinity 11-amino-acid peptide; the data clearly show significant peptide-protein interactions at 6 of the 11 positions of the phosphopeptide, from −2 to +3, relative to the tyrosine residue (225). These are the residues that have associated high electron density, indicating a fixed position in the crystal (except for the side chain portion of Gln-1). In addition to the phosphotyrosine-binding interactions described above, there are several ring interactions that define the rest of the phosphotyrosine pocket. There is also a very well-defined interaction of isoleucine at +3 with a deep pocket in the SH2 domain that results in protection of 95% of the surface of the amino acid side chain. The two glutamate residues at +1 and +2 are on the surface of the protein and largely exposed to solvent. Glu+1 appears to interact through its carboxyl group with a lysine amino group, and Glu+2 appears to be stabilized by a nearby arginine guanidinium and its associated H_2O molecules. The amino acids at positions −1 and −2 appear to cap the phosphotyrosine binding through the polypeptide backbone at position −1 and the proline ring at −2.

Other SH2 domain proteins bind different peptides through interactions at the same +1 to +3 positions relative to the phosphotyrosine. This has been elegantly investigated by Songyang et al. (205) through a study of selectivity of binding of random peptides to different SH2 domains. Although the results obtained in this experiment represent bulk selectivity for certain amino acids at certain positions relative to phosphotyrosine, rather than selectivity of individual peptides of known sequence, the results are clear. Each of the three positions following the phosphotyrosine plays an important role in determining the selectivity of binding in certain SH2 proteins, but the amino acids that are crucial and the extent to which they are crucial differs markedly. Thus, most of the discrimination of the C-terminal SH2 domain of p85 is due to its preference for methionine at +3, whereas most of the discrimination of Nck is at positions 1 and 2, where it prefers glutamate and aspartate, respectively (205).

SH3 Domain

The SH3 domain is a second noncatalytic domain of Src which is involved in protein-protein interactions and which is part of a motif shared by other proteins, including tyrosine kinases, phospholipase C-γ (PLC-γ) PI3K, GTPase-activating protein, the cell proliferation proteins Crk and Grb2/Sem5, and the cytoskeletal proteins spectrin, myosin 1, and an actin-binding protein (see references 117, 120, 154, and 168 for a recent list). More than 27 proteins have been shown to have an SH3 domain, which varies between about 55 and 75 amino acids, and its structure has been determined from four different specific domains: spectrin (154), Src (245), PI3K (120), and PLC (118). Each such structure is composed of antiparallel sheets oriented more or less at right angles to one another (or, for PLC, two partial greek key motifs of a barrel oriented such that the strands on opposite sides cross almost perpendicularly), and the amino acids in the conserved strands and a conserved C-terminal 3_{10} helix correspond to many of those that are conserved among SH3 proteins. In each case, a hydrophobic pocket is formed on the surface of the molecule; those of PI3K and Src are remarkably similar (120), and the location of the pocket is conserved between PLC and spectrin (118). This hydrophobic pocket has been implicated in peptide binding for Src (245), since binding of such a peptide perturbs the signal from these amino acids. There are notable differences among the protein structures; PLC, for example, is very similar in secondary structure to spectrin but not to Src, leading to different architectures (118). This property presumably leads to different binding specificities.

The substrates to which SH3-containing proteins bind include an uncharacterized protein similar to GTPase-activating protein-rho, detected with Abl (36); mSos1 and hSos1 (proteins similar to *Drosophila* Sos, which is required for Ras signaling), detected with Grb2 (187); formin and the rat m4 muscarinic receptor, detected with Abl (181); PI3K, detected with v-Src (130); and $p56^{lck}$ and $p59^{fyn}$ (172, 173).

Like the SH2 domain, the SH3 domain binds simple peptides with a high degree of sequence specificity and a high affinity. As judged on a qualitative basis, a 10-amino-acid proline-rich sequence is responsible for strong binding of the Abl SH3 domain to two proteins, called 3BP-1 and 3BP-2 (36, 181). This binding is specific in two ways. First, some but not all single-amino-acid alterations destroy detectable binding. Thus, prolines at positions 2, 7, and 10 are important but those at 5 and to some extent 9 are not. Nonproline residues do not

appear to be as important, except perhaps at position 1 (181). Second, peptide binding is SH3 domain specific. Thus, 3BP1 binds the SH3 domains of Abl and Src but not those of Neural Src or Crk (36), and 3BP2 binds most strongly to Abl SH3, less so to Src SH3 and Grb2, and poorly to Nck (181).

Similarly, binding of mSos1 to Grb2 appears to be through a proline-rich motif at the C terminus of the protein (187); any of several proline-rich 11-amino-acid peptides corresponding to sequences in this region all compete, and competition appears to require a C-terminal arginine. This arginine may add selectivity to the binding of mSos1 to Grb2. A peptide containing the relevant arginine-containing motif binds to Grb2 through its SH3 domain with a K_d of 25 nM (128).

CONCLUDING REMARKS

Alberts and Miake-Lye (5), summarizing a meeting entitled Proteins as Machines, described Tom Pollard's flow diagram for the detailed analysis of a cell biology process. First, a complete inventory of all the molecules making up the machine must be made. Second, a determination must be made of how and in what order the molecules interact with each other. Third, both detailed rate constants for each transition and structures of each component at atomic resolution must be obtained. While no process is yet completely understood at the three levels described by Pollard, enormous progress has been made in deciphering protein machines. In this review, we have tried to convey some of the classical and more recent approaches used to develop the inventory of proteins and the nature of their interactions.

Two factors are having a large impact on how cellular processes are viewed. First, the vast amount of DNA sequence information being obtained means that the identity of almost all proteins, at the level of primary sequence, may soon be known. Complete sequences for organisms such as *E. coli*, yeast cells, and the nematodes and nearly complete compilations of the cDNA sequences for human tissues should be available in the next few years. Second, the range of new procedures now available means that hundreds to thousands of new protein-protein interactions may be identified in the same period. Ten to twenty years ago, only a few complexes of proteins were well characterized as to their subunit composition and specific interactions; currently, a large number of such complexes are known. Relatively soon, there may be an enormous number. The continuing challenge will be for biochemists and cell, molecular, and structural biologists to use this information to understand how the cell works.

ACKNOWLEDGMENTS

We thank R. Angerer, M. Dumont, E. Grayhack, T. Miller, F. Sherman, and L. Tabak for many helpful comments on the manuscript and J. Nimmo for superb secretarial assistance.

Research from our laboratories has been supported by grants MCB-9118571 from the National Science Foundation and VM-99 from the American Cancer Society (to E.M.P.) and grants GM 49065 and CA 28146 from the National Institutes of Health (to S.F.).

REFERENCES

1. **Ackers, G. K.** 1970. Analytical gel chromatography of proteins. Adv. Protein Chem. 24:343–441.
2. **Ackers, G. K.** 1975. Molecular sieve methods of analysis, p. 1–94. *In* H. Neurath and R. L. Hill (ed.), The proteins. Academic Press, Inc., New York.
3. **Adams, A. E. M., and D. Botstein.** 1989. Dominant suppressors of yeast actin mutations that are reciprocally suppressed. Genetics 121:675–683.
4. **Alber, T.** 1992. Structure of the leucine zipper. Curr. Opin. Genet. Dev. 2:205–210.
5. **Alberts, B., and R. Miake-Lye.** 1992. Unscrambling the puzzle of biological machines: the importance of the details. Cell 68:415–420.
6. **Ayer, D. E., L. Kretzner, and R. N. Eisenman.** 1993. Mad: a heterodimeric partner for Max that antagonizes Myc transcriptional activity. Cell 72:211–222.
7. **Baird, B. A., and G. G. Hammes.** 1976. Chemical cross-linking studies of chloroplast coupling factor 1. J. Biol. Chem. 251:6953–6962.
8. **Ballinger, D. G., N. Xue, and K. D. Harshman.** 1993. A *Drosophila* photoreceptor cell-specific protein, calphotin, binds calcium and contains a leucine zipper. Proc. Natl. Acad. Sci. USA 90:1536–1540.
9. **Banik, U., N. C. Mandal, B. Bhattacharyya, and S. Roy.** 1993. A fluorescence anisotropy study of tetramer-dimer equilibrium of λ repressor and its implication for function. J. Biol. Chem. 268:3938–3943.
10. **Barbas, C. F., A. S. Kang, R. A. Lerner, and S. J. Benkovic.** 1991. Assembly of combinatorial antibody libraries on phage surfaces: the gene III site. Proc. Natl. Acad. Sci. USA 88:7978–7982.
11. **Bass, S., R. Greene, and J. A. Wells.** 1990. Hormone phage: an enrichment method for variant proteins with altered binding properties. Proteins 8:309–314.
12. **Basson, M. E., R. L. Moore, J. O'Rear, and J. Rine.** 1987. Identifying mutations in duplicated functions in *Saccharomyces cerevisiae* by use of the *LYS2* gene. Genetics 117:645–655.
13. **Beckett, D., K. S. Koblan, and G. K. Ackers.** 1991. Quantitative study of protein association at picomolar concentrations: the λ phage cI repressor. Anal. Biochem. 196:69–75.
14. **Beeckmans, S., and L. Kanarek.** 1981. Demonstration of physical interactions between consecutive enzymes of the citric acid cycle and of the aspartate-malate shuttle. Eur. J. Biochem. 117:527–535.
15. **Bender, T., and J. R. Pringle.** 1991. Use of a screen for synthetic lethal and multicopy suppressee mutants to identify two new genes involved in morphogenesis in *Saccharomyces cerevisiae*. Mol. Cell. Biol. 11:1295–1305.
16. **Birge, R. B., J. E. Fajardo, C. Reichman, S. E. Shoelson, Z. Songyang, L. C. Cantley, and H. Hanafusa.** 1993. Identification and characterization of high-affinity interaction between v-Crk and tyrosine-phosphorylated paxillin in CT10-transformed fibroblasts. Mol. Cell. Biol. 13:4648–4656.
17. **Blackwood, E. M., and R. N. Eisenman.** 1991. Max: a helix-loop-helix zipper protein that forms a sequence-specific DNA-binding complex with myc. Science 251:1211–1217.
18. **Blanar, M. A., and W. J. Rutter.** 1992. Interaction cloning: Identification of a helix-loop-helix zipper protein that interacts with c-Fos. Science 256: 1014–1018.
19. **Bleil, J. D., and P. M. Wassarman.** 1990. Identification of a ZP3-binding protein on acrosome-intact mouse sperm by photoaffinity crosslinking. Proc. Natl. Acad. Sci. USA 87:5563–5567.
20. **Booker, G. W., A. L. Breeze, A. K. Downing, G. Panayotou, I. Gout, M. D. Waterfield, and I. D. Campbell.** 1992. Structure of an SH2 domain of the p85α subunit of phosphatidylinositol-3-OH kinase. Nature (London) 358: 684–687.
21. **Bragg, P. D., and C. Hou.** 1980. A cross-linking study of the Ca^{2+}, Mg^{2+}-activated adenosine triphosphatase of *Escherichia coli*. Eur. J. Biochem. 106:495–503.
22. **Bragg, P. D., and C. Hou.** 1986. Chemical crosslinking of α subunits in the F_1 adenosine triphosphatase of *Escherichia coli*. Arch. Biochem. Biophys. 244:361–372.
23. **Brent, R., and M. Ptashne.** 1985. A eukaryotic transcriptional activator bearing the DNA specificity of a prokaryotic repressor. Cell 43:729–736.
24. **Bretscher, A., and K. Weber.** 1980. Villin is a major protein of the microvillus cytoskeleton which binds both G and F actin in a calcium-dependent manner. Cell 20:839–847.
25. **Brown, R. L.** 1992. Functional regions of the inhibitory subunit of retinal rod cGMP phosphodiesterase identified by site-specific mutagenesis and fluorescence spectroscopy. Biochemistry 31:5918–5925.
26. **Burton, D. R., C. F. Barbas, M. A. A. Persson, S. Koenig, R. M. Chanock, and R. A. Lerner.** 1991. A large array of human monoclonal antibodies to type 1 human immunodeficiency virus from combinatorial libraries of asymptomatic seropositive individuals. Proc. Natl. Acad. Sci. USA 88:10134–10137.
27. **Calakos, N., M. K. Bennett, K. E. Peterson, and R. H. Scheller.** 1994. Protein-protein interactions contributing to the specificity of intracellular vesicular trafficking. Science 263:1146–1149.
28. **Carr, C., J. Petersen, P. S. Kim, and T. Alber.** 1992. Unpublished data.
29. **Carr, D. W., and J. D. Scott.** 1992. Blotting and band-shifting: techniques for studying protein-protein interactions. Trends Biochem. Sci. 17:246–249.
30. **Chardin, P., J. H. Camonis, N. W. Gale, L. V. Aelst, J. Schlessinger, M. H. Wigler, and D. Bar-Sagi.** 1993. Human Sos1: a guanine nucleotide exchange factor for Ras that binds to GRB2. Science 260:1338–1343.
31. **Chellappan, S., V. B. Kraus, B. Kroger, K. Munger, P. M. Howley, W. C. Phelps, and J. R. Nevins.** 1992. Adenovirus E1A, simian virus 40 tumor antigen, and human papillomavirus E7 protein share the capacity to disrupt the interaction between transcription factor E2F and the retinoblastoma gene product. Proc. Natl. Acad. Sci. USA 89:4549–4553.
32. **Chellappan, S. P., S. Hiebert, M. Mudryj, J. M. Horowitz, and J. R. Nevins.**

1991. The E2F transcription factor is a cellular target for the RB protein. Cell 65:1053–1061.
33. **Chen, Y., Y. W. Ebright, and R. H. Ebright.** 1994. Identification of the target of a transcription activator protein by protein-protein photocrosslinking. Science 265:90–92.
34. **Chevray, P. M., and D. Nathans.** 1992. Protein interaction cloning in yeast: identification of mammalian proteins that react with the leucine zipper of Jun. Proc. Natl. Acad. Sci. USA 89:5789–5793.
35. **Chien, C.-T., P. L. Bartel, R. Sternglanz, and S. Fields.** 1991. The two-hybrid system: a method to identify and clone genes for proteins that interact with a protein of interest. Proc. Natl. Acad. Sci. USA 88:9578–9582.
36. **Cicchetti, P., B. J. Mayer, G. Thiel, and D. Baltimore.** 1992. Identification of a protein that binds to the SH3 region of Abl and is similar to Bcr and GAP-rho. Science 257:803–806.
37. **Clackson, T., H. R. Hoogenboom, A. D. Griffiths, and G. Winter.** 1991. Making antibody fragments using phage display libraries. Nature (London) 352:624–628.
38. **Cohen, B. D., D. R. Lowy, and J. T. Schiller.** 1993. The conserved C-terminal domain of the bovine papillomavirus E5 oncoprotein can associate with an α-adaptin-like molecule: a possible link between growth factor receptors and viral transformation. Mol. Cell. Biol. 13:6462–6468.
39. **Cornish, V. W., D. R. Benson, C. A. Altenbach, K. Hideg, W. L. Hubbell, and P. G. Schultz.** 1994. Site-specific incorporation of biophysical probes into proteins. Proc. Natl. Acad. Sci. USA 91:2910–2914.
40. **Costigan, C., S. Gehrung, and M. Snyder.** 1992. A synthetic lethal screen identifies SLK1, a novel protein kinase homolog implicated in yeast cell morphogenesis and cell growth. Mol. Cell. Biol. 12:1162–1178.
41. **Cover, J. A., J. M. Lambert, C. M. Norman, and R. R. Traut.** 1981. Identification of proteins at the subunit interface of the *Escherichia coli* ribosome by cross-linking with dimethyl 3,3'-dithiobis(propionimidate). Biochemistry 20:2843–2852.
42. **Crick, F. H. C.** 1953. The packing of α-helices: simple coiled-coils. Acta Crystallogr. 6:689–697.
43. **Cull, M. G., J. F. Miller, and P. J. Schatz.** 1992. Screening for receptor ligands using large libraries of peptides linked to the C terminus of the *lac* repressor. Proc. Natl. Acad. Sci. USA 89:1865–1869.
44. **Cwirla, S. E., E. A. Peters, R. W. Barrett, and W. J. Dower.** 1990. Peptides on phage: a vast library of peptides for identifying ligands. Proc. Natl. Acad. Sci. USA 87:6378–6382.
45. **Dalton, S., and R. Treisman.** 1992. Characterization of SAP-1, a protein recruited by serum response factor to the c-fos serum response element. Cell 68:597–612.
46. **Dang, C. V., J. Barrett, M. Villa-Garcia, L. M. S. Resar, G. J. Kato, and E. R. Fearon.** 1991. Intracellular leucine zipper interactions suggest c-Myc hetero-oligomerization. Mol. Cell. Biol. 11:954–962.
47. **Darawshe, S., G. Rivas, and A. P. Minton.** 1993. Sedimentation equilibrium-quantitative polyacrylamide gel electrophoresis (SE-QPAGE): a new technique for the detection of associations in multicomponent solutions. Anal. Biochem. 215:236–242.
48. **de Gunzberg, J., R. Riehl, and R. A. Weinberg.** 1989. Identification of a protein associated with $p21^{ras}$ by chemical crosslinking. Proc. Natl. Acad. Sci. USA 86:4007–4011.
49. **Denny, J. B., and B. Blobel.** 1984. ^{125}I-labeled crosslinking reagent that is hydrophilic, photoactivatable, and cleavable through an azo linkage. Proc. Natl. Acad. Sci. USA 81:5286–5290.
50. **Devlin, J. J., L. C. Panganiban, and P. E. Devlin.** 1990. Random peptide libraries: a source of specific protein binding molecules. Science 249:404–406.
51. **Diamond, M. I., J. N. Miner, S. K. Yoshinaga, and K. R. Yamamoto.** 1990. Transcription factor interactions: selectors of positive or negative regulation from a single DNA element. Science 249:1266–1272.
52. **Dobzhansky, T.** 1946. Genetics of natural populations. XIII. Recombination and variability in populations of *Drosophila pseudoobscura*. Genetics 31:269–290.
53. **Drubin, D. G., K. G. Miller, and D. Botstein.** 1988. Yeast actin-binding proteins: evidence for a role in morphogenesis. J. Cell Biol. 107:2551–2561.
54. **Dumont, M. E., A. F. Corin, and G. A. Campbell.** 1994. Noncovalent binding of heme induces a compact apocytochrome *c* structure. Biochemistry 33:7368–7378.
55. **Durfee, T., K. Becherer, R.-L. Chen, S. H. Yeh, Y. Yang, A. E. Kilburn, W. H. Lee, and S. J. Elledge.** 1993. The retinoblastoma protein associates with the protein phosphatase type 1 catalytic subunit. Genes Dev. 7:555–569.
56. **Durfee, T., B. Gorovits, C. Hensey, and W.-H. Lee.** *In vivo* and *in vitro* determined binding affinities of complexes involving the retinoblastoma protein exhibit an exponential relationship. Submitted for publication.
57. **Eisenstein, E., and H. K. Schachman.** 1989. Determining the roles of subunits in protein function, p. 135–175. *In* T. E. Creighton (ed.), Protein function: a practical approach. IRL Press, Oxford.
58. **El-Deiry, W. S., T. Tokino, V. E. Velculescu, D. B. Levy, R. Parson, J. M. Trent, D. Lin, W. E. Mercer, K. W. Kinzler, and B. Vogelstein.** 1993. *WAF1*, a potential mediator of p53 tumor suppression. Cell 75:817–825.
59. **Ellenberger, T. E., C. J. Brandl, K. Struhl, and S. C. Harrison.** 1992. The GCN4 basic region leucine zipper binds DNA as a dimer of uninterrupted α helices: crystal structure of the protein-DNA complex. Cell 71:1226–1237.
60. **Evans, P. R., G. W. Farrants, and P. J. Hudson.** 1981. Phosphofructokinase: structure and control. Philos. Trans. R. Soc. London 293:53–62.
61. **Ewen, M. E., B. Faha, E. Harlow, and D. M. Livingston.** 1992. Interaction of p107 with cyclin A independent of complex formation with viral oncoproteins. Science 255:85–87.
62. **Faha, B., M. E. Ewen, L.-H. Tsai, D. M. Livingston, and E. Harlow.** 1992. Interaction between human cyclin A and adenovirus E1A-associated p107 protein. Science 255:87–90.
63. **Fahien, L. A., and S. E. Smith.** 1974. The enzyme-enzyme complex of transaminase and glutamate dehydrogenase. J. Biol. Chem. 249:2696–2703.
64. **Felder, S., M. Zhou, P. Hu, J. Urena, A. Ullrich, M. Chaudhuri, S. E. Shoelson, and J. Schlessinger.** 1993. SH2 domains exhibit high-affinity binding to tyrosine-phosphorylated peptides yet also exhibit rapid dissociation and exchange. Mol. Cell. Biol. 13:1449–1455.
65. **Fields, S., and O.-K. Song.** 1989. A novel genetic system to detect protein-protein interactions. Nature (London) 340:245–246.
66. **Fields, S., and R. Sternglanz.** 1994. The two-hybrid system: An assay for protein-protein interactions. Trends Genet. 10:286–292.
67. **Finlay, C. A., P. W. Hinds, and A. J. Levine.** 1989. The p53 proto-oncogene can act as a suppressor of transformation. Cell 57:1083–1093.
68. **Flynn, D. C., T.-H. Leu, A. B. Reynolds, and J. T. Parson.** 1993. Identification and sequence analysis of cDNAs encoding a 110-kilodalton actin filament-associated $pp60^{src}$ substrate. Mol. Cell. Biol. 13:7892–7900.
69. **Formosa, T., J. Barry, B. M. Alberts, and J. Greenblatt.** 1991. Using protein affinity chromatography to probe structure of protein machines. Methods Enzymol. 208:24–45.
70. **Friedrich, P.** 1984. Supramolecular enzyme organization: quarternary structure and beyond. Pergamon Press, Oxford.
71. **Friguet, B., A. F. Chaffotte, L. Djavadi-Ohaniance, and M. Goldberg.** 1985. Measurements of the true affinity constant in solution of antigen-antibody complexes. B1 enzyme-linked immunosorbent assay. J. Immunol. Methods 77:305–319.
72. **Gegner, J. A., and F. W. Dahlquist.** 1991. Signal transduction in bacteria: CheW forms a reversible complex with the protein kinase CheA. Proc. Natl. Acad. Sci. USA 88:750–754.
73. **Geiser, J. R., H. A. Sundberg, B. H. Chang, E. G. D. Muller, and T. N. Davis.** 1993. The essential mitotic target of calmodulin is the 110-kilodalton component of the spindle pole body in *Saccharomyces cerevisiae*. Mol. Cell. Biol. 13:7913–7924.
74. **Germino, F. J., Z. X. Wang, and S. M. Weissman.** 1993. Screening for *in vivo* protein-protein interactions. Proc. Natl. Acad. Sci. USA 90:933–937.
75. **Gilbert, G. A., and G. L. Kellett.** 1971. Interacting systems of the type A+B=C: ovalbumin and myoclobin. J. Biol. Chem. 246:6079–6086.
76. **Gill, S. C., S. E. Weitzel, and P. H. Von Hippel.** 1991. *Escherichia coli* σ^{70} and NusA proteins. 1. Binding interactions with core RNA polymerase in solution and within the transcription complex. J. Mol. Biol. 220:307–324.
77. **Glenney, J. R., and K. Weber.** 1993. Detection of calmodulin-binding polypeptides separated in SDS-polyacrylamide gels by a sensitive [^{125}I]calmodulin gel overlay assay. Methods Enzymol. 102:204–210.
78. **Goodsell, D. S., and A. J. Olson.** 1993. Soluble proteins: size, shape and function. Trends Biochem. Sci. 18:65–68.
79. **Gorlich, D., E. Hartmann, S. Prehn, and T. A. Rapoport.** 1992. A protein of the endoplasmic reticulum involved early in polypeptide translocation. Nature (London) 357:47–52.
80. **Greenblatt, J., and J. Li.** 1981. The *nusA* gene protein of *Escherichia coli*. J. Mol. Biol. 147:11–23.
81. **Gross, M., and T. Wallimann.** 1993. Kinetics of assembly and dissociation of the mitochondrial creatine kinase octamer: a fluorescence study. Biochemistry 32:13933–13940.
82. **Hadwiger, J. A., C. Wittenberg, H. E. Richardson, M. D. Lopes, and S. I. Reed.** 1989. A family of cyclin homologs that control the G1 phase in yeast. Proc. Natl. Acad. Sci. USA 86:6255–6259.
83. **Hannon, G. J., D. Demetrick, and D. Beach.** 1993. Isolation of the Rb-related p130 through its interaction with CDK2 and cyclins. Genes Dev. 7:2378–2391.
84. **Harlow, E., P. Whyte, B. R. Franza, and C. Schley.** 1986. Association of adenovirus early-region 1A proteins with cellular polypeptides. Mol. Cell. Biol. 6:1579–1589.
85. **Harper, J. W., G. R. Adami, N. Wei, K. Keyomarsi, and S. J. Elledge.** 1993. The p21 Cdk-interacting protein Cip1 is a potent inhibitor of G1 cyclin-dependent kinases. Cell 75:805–816.
86. **Hartman, P. E., and J. R. Roth.** 1973. Mechanisms of suppression. Adv. Genet. 17:1–105.
87. **Hartmann, E., M. Wiedmann, and T. A. Rapoport.** 1989. A membrane component of the endoplasmic reticulum that may be essential for protein translocation. EMBO J. 8:2225–2229.
88. **Herberg, F. W., W. R. G. Dostmann, M. Zorn, S. J. Davis, and S. S. Taylor.** 1994. Crosstalk between domains in the regulatory subunit of cAMP-dependent protein kinase: Influence of amino terminus on cAMP binding and

holoenzyme formation. Biochemistry 33:7485–7494.
89. **Herberg, F. W., and S. S. Taylor.** 1993. Physiological inhibitors of the catalytic subunit of cAMP-dependent protein kinase: effect of MgATP on protein-protein interactions. Biochemistry 32:14015–14022.
90. **Herskowitz, I.** 1987. Functional inactivation of genes by dominant negative mutations. Nature (London) 329:219–222.
91. **Heyduk, T., J. C. Lee, Y. W. Ebright, E. E. Blatter, Y. Zhou, and R. H. Ebright.** 1993. CAP interacts with RNA polymerase in solution in the absence of promoter DNA. Nature (London) 364:548–549.
92. **Hiebert, S. W., S. P. Chellappan, J. M. Horowitz, and J. R. Nevins.** 1992. The interaction of RB with E2F coincides with an inhibition of the transcriptional activity of E2F. Genes Dev. 6:177–185.
93. **Hieter, P., C. Mann, M. Snyder, and R. W. Davis.** 1985. Mitotic stability of yeast chromosomes: a colony color assay that measures nondisjunction and chromosome loss. Cell 40:381–392.
94. **High, S., B. Bartoglio, D. Gorlich, S. S. L. Andersen, A. J. Ashford, A. Giner, E. Hartmann, S. Prehn, T. A. Rapoport, B. Dobberstein, and J. Brunner.** 1993. Site-specific photocross-linking reveals that Sec61p and TRAM contact different regions of a membrane-inserted signal sequence. J. Biol. Chem. 368:28745–28751.
95. **Hill, R. L., and K. Brew.** 1975. Lactose synthetase. Adv. Enzymol. 43:411–490.
96. **Hodgkin, J., K. Kondo, and R. H. Waterston.** 1987. Suppression in the nematode *Caenorhabditis elegans*. Trends Genet. 3:325–329.
97. **Hope, I. A., and K. Struhl.** 1986. Functional dissection of a eukaryotic transcriptional activator protein, GCN4 of yeast. Cell 46:885–894.
98. **Hu, J. C., E. K. O'Shea, P. S. Kim, and R. T. Sauer.** 1990. Sequence requirements for coiled-coils: analysis with λ repressor-GCN4 leucine zipper fusions. Science 250:1400–1403.
99. **Hu, P., B. Margolis, E. Y. Skolnik, R. Lammers, A. Ullrich, and J. Schlessinger.** 1992. Interaction of phosphatidylinositol 3-kinase-associated p85 with epidermal growth factor and platelet-derived growth factor receptors. Mol. Cell. Biol. 12:981–990.
100. **Huang, S., W.-H. Lee, and E. Y.-H. P. Lee.** 1991. A cellular protein that competes with SV40 T antigen for binding to the retinoblastoma gene product. Nature (London) 350:160–162.
101. **Huffaker, T. C., M. A. Hoyt, and D. Botstein.** 1987. Genetic analysis of the yeast cytoskeleton. Annu. Rev. Genet. 21:259–284.
102. **Hummel, J. P., and W. J. Dreyer.** 1962. Measurement of protein-binding phenomena by gel filtration. Biochim. Biophys. Acta 63:530–532.
103. **Hurley, J. B., and L. Stryer.** 1992. Purification and characterization of the γ regulatory subunit of the cyclic GMP phosphodiesterase from retinal rod outer segments. J. Biol. Chem. 257:11094–11099.
104. **Il'ichev, A. A., O. O. Minenkova, S. I. Tatkov, N. N. Karpyshev, A. M. Eroshkin, V. A. Petrenko, and L. S. Sandakhchiev.** 1989. M13 filamentous bacteriophage in protein engineering. Dokl. Akad. Nauk SSSR 307:431–433.
105. **Iwabuchi, K., B. Li, P. L. Bartel, and S. Fields.** 1993. Use of the two-hybrid system to identify the domain of p53 involved in oligomerization. Oncogene 8:1693–1696.
106. **James, P., M. Inui, M. Tada, M. Chiesi, and E. Carafoli.** 1989. Nature and site of phospholamban regulation of the Ca^{2+} pump of sarcoplasmic reticulum. Nature (London) 342:90–92.
107. **Jarvik, J., and D. Botstein.** 1975. Conditional-lethal mutations that suppress genetic defects in morphogenesis by altering structural proteins. Proc. Natl. Acad. Sci. USA 72:2738–2742.
108. **Jarvis, T. C., D. M. Ring, S. S. Daube, and P. H. Von Hippel.** 1990. "Macromolecular crowding": thermodynamic consequences for protein-protein interactions within the T4 DNA replication complex. J. Biol. Chem. 265:15160–15167.
109. **Jonsson, U., L. Fagerstam, B. Ivarsson, B. Johnsson, R. Karlsson, K. Lundh, S. Lofas, B. Persson, H. Roos, I. Ronnberg, S. Sjolander, E. Stenberg, R. Stahlberg, C. Urbaniczky, H. Ostlin, and M. Malmqvist.** 1991. Real-time biospecific interaction analysis using surface plasmon resonance and a sensor chip technology. BioTechniques. 11:620–627.
110. **Kang, A. S., C. F. Barbas, K. D. Janda, S. J. Benkovic, and R. A. Lerner.** 1991. Linkage of recognition and replication functions by assembling combinatorial antibody Fab libraries along phage surfaces. Proc. Natl. Acad. Sci. USA 88:4363–4366.
111. **Karlsson, R., A. Michaelsson, and L. Mattsson.** 1991. Kinetic analysis of monoclonal antibody-antigen interactions with a new biosensor based analytical system. J. Immunol. Methods 145:229–240.
112. **Keegan, L., G. Gill, and M. Ptashne.** 1986. Separation of DNA binding from the transcription-activating function of a eukaryotic regulatory protein. Science 231:699–704.
113. **Kellogg, D. R., C. M. Field, and B. M. Alberts.** 1989. Identification of microtubule-associated proteins in the centrosome, spindle, and kinetochore of the early *Drosophila* embryo. J. Cell Biol. 109:2977–2991.
114. **Kerppola, T. K., and T. Curran.** 1991. Fos-Jun heterodimers and Jun homodimers bend DNA in opposite orientations: implications for transcription factor cooperativity. Cell 66:317–326.
115. **Kim, Y. T., S. Tabor, J. E. Churchich, and C. C. Richardson.** 1992. Interactions of gene 2.5 protein and DNA polymerase of bacteriophage T7. J. Biol. Chem. 267:15032–15040.
116. **Klotz, I. M., D. W. Darnall, and N. R. Langerman.** 1975. Quarternary structure of proteins, p. 293–411. *In* H. Neurath and R. L. Hill (ed.), The proteins. Academic Press, Inc., New York.
117. **Koch, C. A., D. Anderson, M. F. Moran, C. Willis, and T. Pawson.** 1991. SH2 and SH3 domains: elements that control interactions of cytoplasmic signaling proteins. Science 252:668–674.
118. **Kohda, D., H. Hatanaka, M. Odaka, V. Mandiyan, A. Ullrich, J. Schlessinger, and F. I. Nagaki.** 1993. Solution structure of the SH3 domain of phospholipase C-g. Cell 72:953–960.
119. **Koshland, D., J. C. Kent, and L. H. Hartwell.** 1985. Genetic analysis of the mitotic transmission of minochromosomes. Cell 40:393–403.
120. **Koyama, S., H. Yu, D. C. Dalgarno, T. B. Shin, L. D. Zydowsky, and S. L. Schreiber.** 1993. Structure of the PI3K SH3 domain and analysis of the SH3 family. Cell 72:945–952.
121. **Krieg, U. C., A. E. Johnson, and P. Walter.** 1989. Protein translocation across the endoplasmic reticulum membrane: identification by photocross-linking of a 39-kD integral membrane glycoprotein as part of a putative translocation tunnel. J. Cell Biol. 109:2033–2043.
122. **Kukuruzinska, M. A., B. W. Turner, G. K. Ackers, and S. Roseman.** 1984. Subunit association of enzyme I of the *Salmonella typhimurium* phosphoenolpyruvate:glycose phosphotransferase system. J. Biol. Chem. 259:11679–11681.
123. **Kwon, O.-S., and J. E. Churchich.** 1994. Interaction of 70-kDa heat shock cognate protein with peptides and *myo*-inositol monophosphatase. J. Biol. Chem. 269:266–271.
124. **Landschulz, W. H., P. F. Johnson, and S. L. McKnight.** 1988. The leucine zipper: a hypothetical structure common to a new class of DNA binding proteins. Science 240:1759–1764.
125. **Lazarides, E., and U. Lindberg.** 1974. Actin is the naturally occurring inhibitor of deoxyribonuclease I. Proc. Natl. Acad. Sci. USA 71:4742–4746.
126. **Lee, F. S., D. S. Auld, and B. L. Vallee.** 1989. Tryptophan fluorescence as a probe of placental ribonuclease inhibitor binding to angiogenin. Biochemistry 28:219–224.
126a. **Lee, F. S., R. Shapiro, and B. L. Vallee.** 1989. Tight-binding inhibition of angiogenin and ribonuclease A by placental ribonuclease inhibitor. Biochemistry 28:225–230.
127. **Li, B., and S. Fields.** 1993. Identification of mutations in p53 that affect its binding to SV40 large T antigen by using the yeast two-hybrid system. FASEB J. 7:957–963.
128. **Li, N., A. Batzer, R. Daly, V. Yajnik, E. Skolnik, P. Chardin, D. Bar-Sagi, B. Margolis, and J. Schlessinger.** 1993. Guanine-nucleotide-releasing factor hSos1 binds to Grb2 and links receptor tyrosine kinases to Ras signalling. Nature (London) 363:85–88.
129. **Li, S., and J. M. Sedivy.** 1993. Raf-1 protein kinase activates the NF-κB transcription factor by dissociating the cytoplasmic NF-kB-IkB complex. Proc. Natl. Acad. Sci. USA 90:9247–9251.
130. **Liu, X., L. E. M. Marengere, C. A. Koch, and T. Pawson.** 1993. The v-Src SH3 domain binds phosphatidylinositol 3'-kinase. Mol. Cell. Biol. 13:5225–5232.
131. **Lowman, H. B., S. H. Bass, N. Simpson, and J. A. Wells.** 1991. Selecting high-affinity binding proteins by monovalent phage display. Biochemistry 30:10832–10838.
132. **Ludlow, J. W., J. A. DeCaprio, C.-M. Huang, W.-H. Lee, E. Paucha, and D. M. Livingston.** 1989. SV40 large T antigen binds preferentially to an underphosphorylated member of the retinoblastoma susceptibility gene product family. Cell 56:57–65.
133. **Ludlow, J. W., C. L. Glendening, D. M. Livingston, and J. A. DeCaprio.** 1993. Specific enzymatic dephosphorylation of the retinoblastoma protein. Mol. Cell. Biol. 13:367–372.
134. **Ma, J., and M. Ptashne.** 1987. A new class of yeast transcriptional activators. Cell 51:113–119.
135. **MacGregor, P. F., C. Abate, and T. Curran.** 1990. Direct cloning of leucine zipper proteins: jun binds cooperatively to the CRE with CRE-BP1. Oncogene 5:451–458.
136. **Malmqvist, M.** 1993. Biospecific interaction analysis using biosensor technology. Nature (London) 361:186–187.
137. **Marengere, L. E. M., Z. Songyang, G. D. Gish, M. D. Schaller, J. T. Parsons, M. J. Stern, L. C. Cantley, and T. Pawson.** 1994. SH2 domain specificity and activity modified by a single residue. Nature (London) 369:502–505.
138. **Mason, S. W., J. Li, and J. Greenblatt.** 1992. Direct interaction between two *Escherichia coli* transcription antitermination factors, NusB and ribosomal protein S10. J. Mol. Biol. 223:55–66.
139. **Mayer, B. J., P. K. Jackson, and D. Baltimore.** 1991. The noncatalytic src homology region 2 segment of *abl* tyrosine kinase binds to tyrosine-phosphorylated cellular proteins with high affinity. Proc. Natl. Acad. Sci. USA 88:627–631.
140. **Mayer, B. J., P. K. Jackson, R. A. Van Etten, and D. Baltimore.** 1992. Point mutations in the *abl* SH2 domain coordinately impair phosphotyrosine binding in vitro and transforming activity in vivo. Mol. Cell. Biol. 12:609–618.

141. **McCafferty, J., A. D. Griffiths, G. Winter, and D. J. Chiswell.** 1990. Phage antibodies: filamentous phage displaying antibody variable domains. Nature (London) 348:552–554.
142. **Meeks-Wagner, D., and L. H. Hartwell.** 1986. Normal stoichiometry of histone dimer sets is necessary for high fidelity of mitotic chromosome transmission. Cell 44:43–52.
143. **Miller, J. H.** 1978. The *lacI* gene: its role in *lac* operon control and its use as a genetic system, p. 31–88. *In* J. H. Miller and W. S. Reznikoff (ed.), The operon. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
144. **Miller, K. G., and B. M. Alberts.** 1989. F-actin affinity chromatography: technique for isolating previously unidentified actin-binding proteins. Proc. Natl. Acad. Sci. USA 86:4808–4812.
145. **Miller, K. G., C. M. Field, B. M. Alberts, and D. R. Kellogg.** 1991. Use of actin filament and microtubule affinity chromatography to identify proteins that bind to the cytoskeleton. Methods Enzymol. 196:303–319.
146. **Mills, J. S., M. P. Walsh, K. Nemcek, and J. D. Johnson.** 1988. Biologically active fluorescent derivatives of spinach calmodulin that report calmodulin target protein binding. Biochemistry 27:991–996.
147. **Mita, S., A. Tominaga, Y. Hitoshi, K. Sakamoto, T. Honjo, M. Akagi, Y. Kikuchi, N. Yamaguchi, and K. Takatsu.** 1989. Characterization of high-affinity receptors for interleukin 5 on interleukin 5-dependent cell lines. Proc. Natl. Acad. Sci. USA 86:2311–2315.
148. **Mitchell, D. A., T. K. Marshall, and R. J. Deschenes.** 1993. Vectors for the inducible overexpression of glutathione S-transferase fusion proteins in yeast. Yeast 9:715–723.
149. **Mitsuzawa, H., I. Uno, T. Oshima, and T. Ishikawa.** 1989. Isolation and characterization of temperature-sensitive mutations in the *RAS2* and *CYR1* genes of *Saccharomyces cerevisiae*. Genetics 123:739–748.
150. **Moerman, D. G., G. M. Benian, R. J. Barstead, L. A. Schriefer, and R. H. Waterston.** 1988. Identification and intracellular localization of the *unc-22* gene product of *Caenorhabditis elegans*. Genes Dev. 2:93–105.
151. **Moerman, D. G., S. Plurad, and R. H. Waterston.** 1982. Mutations in the *unc-54* myosin heavy chain gene of *Caenorhabditis elegans* that alter contractility but not muscle structure. Cell 29:773–781.
152. **Moir, D., S. E. Stewart, B. C. Osmond, and D. Botstein.** 1982. Cold-sensitive cell-division-cycle mutants of yeast: isolation, properties, and pseudoreversion studies. Genetics 100:547–563.
153. **Monod, J., J. Wyman, and J.-P. Changeux.** 1965. On the nature of allosteric transitions: a plausible model. J. Mol. Biol. 12:88–118.
154. **Musacchio, A., M. Noble, R. Pauptit, R. Wierenga, and M. Saraste.** 1992. Crystal structure of a Src-homology 3 (SH3) domain. Nature (London) 359:851–855.
155. **Musch, A., M. Wiedmann, and T. A. Rapoport.** 1992. Yeast Sec proteins interact with polypeptides traversing the endoplasmic reticulum membrane. Cell 69:343–352.
156. **Nakai, H., and C. C. Richardson.** 1990. The gene 1.2 protein of bacteriophage T7 interacts with the *Escherichia coli* dGTP triphosphohydrolase to form a GTP-binding protein. J. Biol. Chem. 265:4411–4419.
157. **Nefsky, B., and A. Bretscher.** 1992. Yeast actin is relatively well behaved. Eur. J. Biochem. 206:949–955.
158. **Nelson, R. M., and G. L. Long.** 1991. Solution-phase equilibrium binding interaction of human protein S with C4b-binding protein. Biochemistry 30:2384–2390.
159. **Nogi, Y., K. Matsumoto, A. Toh-e, and Y. Oshima.** 1977. Regulation of expression of the galactose gene cluster in *Saccharomyces cerevisiae*. II. The isolation and dosage effect of the regulatory gene Gal80. Mol. Gen. Genet. 195:29–34.
160. **Novick, P., B. C. Osmond, and D. Botstein.** 1989. Suppressors of yeast actin mutations. Genetics 121:659–674.
161. **O'Shea, E. K., J. D. Klemm, P. S. Kim, and T. Alber.** 1991. X-ray structure of the GCN4 leucine zipper, a two-stranded, parallel coiled coil. Science 254:539–544.
162. **O'Shea, E. K., R. Rutkowski, and P. S. Kim.** 1992. Mechanism of specificity in the Fos-Jun oncoprotein heterodimer. Cell 68:699–708.
163. **O'Shea, E. K., R. Rutkowski, W. F. Stafford III, and P. S. Kim.** 1989. Preferential heterodimer formation by isolated leucine zippers from fos and jun. Science 245:646–648.
164. **Otto-Bruc, A., B. Antonny, T. M. Vuong, P. Chardin, and M. Chabre.** 1993. Interaction between retinal cyclic GMP phosphodiesterase inhibitor and transducin: kinetics and affinity studies. Biochemistry 32:8636–8645.
165. **Overduin, M., C. B. Rios, B. J. Mayer, D. Baltimore, and D. Cowburn.** 1992. Three dimensional solution structure of the scr homology 2 domain of c-abl. Cell 70:697–704.
166. **Panayotou, G., G. Gish, P. End, O. Truong, I. Gout, R. Dhand, M. J. Fry, I. Hiles, T. Pawson, and M. D. Waterfield.** 1993. Interactions between SH2 domains and tyrosine-phosphorylated platelet-derived growth factor β-receptor sequences: analysis of kinetic parameters by a novel biosensor-based approach. Mol. Cell. Biol. 13:3567–3576.
167. **Pannekoek, H., M. VanMeijer, R. R. Schleef, D. J. Loskutoff, and C. F. Barbas.** 1993. Functional display of human plasminogen-activator inhibitor 1 (PAI-1) on phages: novel perspectives for structure-function analysis by error-prone DNA synthesis. Gene 128:135–140.
168. **Pawson, T., and G. D. Gish.** 1992. SH2 and SH3 domains: from structure to function. Cell 71:359–362.
169. **Phizicky, E. M., R. C. Schwartz, and J. Abelson.** 1986. *Saccharomyces cerevisiae* tRNA ligase. J. Biol. Chem. 261:2978–2986.
170. **Pinkney, M., and J. G. Hoggett.** 1988. Binding of the cyclic AMP receptor protein of *Escherichia coli* to RNA polymerase. Biochem. J. 250:897–902.
171. **Porpaczy, Z., B. Sumegi, and I. Alkonyi.** 1983. Association between the a-ketoglutarate dehydrogenase complex and succinate thiokinase. Biochim. Biophys. Acta 749:172–179.
172. **Prasad, K. V. S., O. Janssen, R. Kapeller, M. Raab, L. C. Cantley, and C. E. Rudd.** 1993. Src-homology 3 domain of protein kinase $p59^{fyn}$ mediates binding to phosphatidylinositol 3-kinase in T cells. Proc. Natl. Acad. Sci. USA 90:7366–7370.
173. **Prasad, K. V. S., R. Kapeller, O. Janssen, H. Repke, J. S. Duke-Cohan, L. C. Cantley, and C. E. Rudd.** 1993. Phosphatidylinositol (PI) 3-kinase and PI 4-kinase binding to the CD4-p56lck SH3 domain binds to PI 3-kinase but not PI 4-kinase. Mol. Cell Biol. 13:7708–7717.
174. **Prelich, G., C.-K. Tan, M. Kostura, M. B. Mathews, A. G. So, K. M. Downey, and B. Stillman.** 1989. Functional identity of proliferating cell nuclear antigen and a DNA polymerase-δ auxiliary protein. Nature (London) 326:517–520.
175. **Rabidran, S. K., R. I. Haroun, J. Clos, J. Wisniewski, and C. Wu.** 1993. Regulation of heat shock factor trimer formation: role of a conserved leucine zipper. Science 259:230–234.
176. **Raff, E. C., and M. T. Fuller.** 1984. Genetic analysis of microtubule function in *Drosophila*, p. 293–304. *In* G. G. Borisy, D. W. Cleveland, and D. B. Murphy (ed.), Molecular biology of the cytoskeleton. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
177. **Ratner, D.** 1974. The interaction of bacterial and phage proteins with immobilized *Escherichia coli* RNA polymerase. J. Mol. Biol. 88:373–383.
178. **Rauhut, R., P. R. Green, and J. Abelson.** 1990. Yeast tRNA-splicing endonuclease is a heterotrimeric enzyme. J. Biol. Chem. 265:18180–18184.
179. **Ray, S. K., M. Arroyo, S. Bigchi, and P. Raychaudhuri.** 1992. Identification of a 60-kilodalton Rb-binding protein, RMP60, that allows the Rb-E2F complex to bind DNA. Mol. Cell. Biol. 12:4327–4333.
180. **Redl, B., J. Walleczek, M. Stoffler-Meilicke, and G. Stoffler.** 1989. Immunoblotting analysis of protein-protein crosslinks within the 50S ribosomal subunit of *Escherichia coli*: a study using dimethylsuberimidate as crosslinking reagent. Eur. J. Biochem. 181:351–356.
181. **Ren, R., B. J. Mayer, P. Cicchetti, and D. Baltimore.** 1993. Identification of a ten-amino acid proline-rich SH3 binding site. Science 259:1157–1161.
182. **Rine, J.** 1991. Gene overexpression in studies of *Saccharomyces cerevisiae*. Methods Enzymol. 194:239–251.
183. **Rivas, G., K. C. Ingham, and A. P. Minton.** 1992. Calcium-linked self-association of human complement Cls. Biochemistry 31:11707–11712.
184. **Rivas, G., K. C. Ingham, and A. P. Minton.** 1994. Ca^{2+}-linked association of human complement Cls and Clr. Biochemistry 33:2341–2348.
185. **Rivas, G., and A. P. Minton.** 1993. New developments in the study of biomolecular associations via sedimentation equilibrium. Trends Biochem. Sci. 18:284–287.
186. **Roberts, B. L., W. Markland, A. C. Ley, R. B. Kent, D. W. White, S. K. Gutterman, and R. C. Ladner.** 1992. Directed evolution of a protein: selection of potent neutrophil elastase inhibitors displayed on M13 fusion phage. Proc. Natl. Acad. Sci. USA 89:2429–2433.
187. **Rozakis-Adcock, M., R. Fernley, J. Wade, T. Pawson, and D. Bowtell.** 1993. The SH2 and SH3 domains of mammalian Grb2 couple the EGF receptor to the Ras activator mSos1. Nature (London) 363:83–85.
188. **Russell, R. B., J. Breed, and G. J. Barton.** 1992. Conservation analysis and structure prediction of the SH2 family of phosphotyrosine binding domains. FEBS Lett. 304:15–20.
189. **Sadowski, I., J. C. Stone, and T. Pawson.** 1986. A noncatalytic domain conserved among cytoplasmic protein-tyrosine kinases modifies the kinase function and transforming activity of Fujinami sarcoma virus $P130^{gag\text{-}fps}$. Mol. Cell. Biol. 6:4396–4408.
190. **Salminen, A., and P. J. Novick.** 1987. A *ras*-like protein is required for a post-Golgi event in yeast secretion. Cell 49:527–538.
191. **Sanders, S. L., K. M. Whitfield, J. P. Vogel, M. D. Rose, and R. W. Schekman.** 1992. Sec61p and BiP directly facilitate polypeptide translocation into the ER. Cell 69:353–365.
192. **Sarkar, F. H., and S. L. Gupta.** 1984. Receptors for human γ interferon: binding and crosslinking of ^{125}I-labeled recombinant human γ interferon to receptors on WISH cells. Proc. Natl. Acad. Sci. USA 81:5160–5164.
193. **Sauer, R. T.** 1979. Molecular characterization of the repressor and its gene cI. Ph.D. thesis. Harvard University, Cambridge, Mass.
194. **Sawyers, C. L., W. Callahan, and O. N. Witte.** 1992. Dominant negative MYC blocks transformation by ABL oncogenes. Cell 70:901–910.
195. **Scherer, P. E., U. C. Manning-Krieg, P. Jeno, G. Schatz, and M. Horst.** 1992. Identification of a 45-kDa protein at the protein import site of the yeast mitochondrial inner membrane. Proc. Natl. Acad. Sci. USA 89:11930–11934.
196. **Schmidt-Dorr, T., P. Oertel-Buchheit, C. Pernelle, L. Bracco, M. Schnarr, and M. Granger-Schnarr.** 1991. Construction, purification and character-

ization of a hybrid protein comprising the DNA binding domain of the LexA repressor and the jun leucine zipper: a circular dichroism and mutagenesis study. Biochemistry 30:9657–9664.
197. **Schuster, S. C., R. V. Swanson, L. A. Alex, R. B. Bourret, and M. I. Simon.** 1993. Assembly and function of a quarternary signal transduction complex monitored by surface plasmon resonance. Nature (London) 365:343–346.
198. **Scott, J. K., and G. P. Smith.** 1990. Searching for peptide ligands with an epitope library. Science 249:386–390.
199. **Shapiro, R., and B. L. Vallee.** 1991. Interaction of human placental ribonuclease with placental ribonuclease inhibitor. Biochemistry 30:2246–2255.
200. **Sikela, J. M., and W. E. Hahn.** 1987. Screening an expression library with a ligand probe: isolation and sequence of a cDNA corresponding to a brain calmodulin-binding protein. Proc. Natl. Acad. Sci. USA 84:3038–3042.
201. **Skolnik, E. Y., B. Margolis, M. Mohammadi, E. Lowenstein, R. Fischer, A. Drepps, A. Ullrich, and J. Schlessinger.** 1991. Cloning of PI3 kinase-associated p85 utilizing a novel method for expression/cloning of target proteins for receptor tyrosine kinases. Cell 65:83–90.
202. **Smith, D. B., and K. S. Johnson.** 1988. Single-step purification of polypeptides expressed in *Escherichia coli* as fusions with glutathione *S*-transferase. Gene 67:31–40.
203. **Smith, G. P.** 1985. Filamentous fusion phage: novel expression vectors that display cloned antigens on the virion surface. Science 228:1315–1317.
204. **Sollner, T., J. Rassow, M. Wiedmann, J. Schlossmann, P. Keil, W. Neupert, and N. Pfanner.** 1992. Mapping of the protein import machinery in the mitochondrial outer membrane by crosslinking of translocation intermediates. Nature (London) 355:84–87.
205. **Songyang, Z., S. E. Shoelson, M. Chaudhuri, G. Gish, T. Pawson, W. G. Haser, F. King, T. Robertys, S. Ratnofsky, R. J. Lechleider, B. G. Neel, R. B. Birge, J. E. Fajardo, M. M. Chou, H. Hanafusa, B. Schaffhausen, and L. C. Cantley.** 1993. SH2 domains recognize specific phosphopeptide sequences. Cell 72:767–778.
206. **Sopta, M., R. W. Carthew, and J. Greenblatt.** 1985. Isolation of three proteins that bind to mammalian RNA polymerase II. J. Biol. Chem. 260:10353–10360.
207. **Soutar, A. K., and D. P. Wade.** 1990. Ligand blotting, p. 55–76. *In* T. Creighton (ed.), Protein function: a practical approach. IRL Press, Oxford.
208. **Srere, P. A.** 1987. Complexes of sequential metabolic enzymes. Annu. Rev. Biochem. 56:89–124.
209. **Steitz, T. A., W. F. Anderson, R. J. Fletterick, and C. M. Anderson.** 1977. High resolution crystal structures of yeast hexokinase complexes with substrates, activators, and inhibitors. J. Biol. Chem. 252:4494–4500.
210. **Stephen, C. W., and D. P. Lane.** 1992. Mutant conformation of p53: precise epitope mapping using a filamentous phage epitope library. J. Mol. Biol. 225:577–583.
211. **Stevens, F. J.** 1986. Analysis of protein-protein interaction by simulation of small-zone size-exclusive chromatography: application to an antibody-antigen association. Biochemistry 25:981–993.
212. **Stoffler, G., B. Redl, J. Walleczek, and M. Stoffler-Meilicke.** 1988. Identification of protein-protein cross-links within the *Escherichia coli* ribosome by immunobloting techniques. Methods Enzymol. 164:64–76.
213. **Susskind, M. M., and P. Youderian.** 1983. Bacteriophage P22 antirepressor and its control. p. 347–363. *In* R. W. Hendrix, J. W. Roberts, F. W. Stahl, and R. A. Weisberg (ed.), Lambda II. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
214. **Tompa, P., J. Batke, J. Ovadi, G. R. Welch, and P. A. Srere.** 1987. Quantitation of the interaction between citrate synthase and malate dehydrogenase. J. Biol. Chem. 282:6089–6092.
215. **Traut, R. R., C. Casiano, and N. Zecherle.** 1989. Crosslinking of protein subunits and ligands by the introduction of disulphide bonds, p. 101–133. *In* T. E. Creighton (ed.), Protein function: a practical approach. IRL Press, Oxford.
216. **Truant, R., H. Xiao, C. J. Ingles, and J. Greenblatt.** 1993. Direct interaction between the transcriptional activation domain of human p53 and the TATA box-binding protein. J. Biol. Chem. 268:2284–2287.
217. **Uchiumi, T., A. J. Wahba, and R. R. Traut.** 1987. Topography and stoichiometry of acidic proteins in large ribosomal subunits from *Artemia salina* as determined by crosslinking. Proc. Natl. Acad. Sci. USA 84:5580–5584.
218. **Urena, J., et al.** 1993. Unpublished data.
219. **Vallee, R. B., and C. A. Collins.** 1986. Purification of microtubules and microtubule-associated proteins from sea urchin eggs and cultured mammalian cells using Taxol, and use of exogenous taxol-stabilized brain microtubules for purifying microtubule-associated proteins. Methods Enzymol. 134:116–121.
220. **Van Aelst, L., M. Barr, S. Marcus, A. Polverino, and M. Wigler.** 1993. Complex formation between RAS and RAF and other protein kinases. Proc. Natl. Acad. Sci. USA 90:6213–6217.
221. **Vincent, J.-P., and M. Lazdunski.** 1972. Trypsin-pancreatic trypsin inhibitor association. Dynamics of the interaction and role of disulfide bridges. Biochemistry 11:2967–2977.
221a. **Vinson, C. R., T. Hai, and S. M. Boyd.** 1993. Dimerization specificity of the leucine zipper-containing bZIP motif on DNA binding: prediction and rational design. Genes Dev. 7:1047–1058.
222. **Vinson, C. R., K. L. LaMarco, P. F. Johnson, W. H. Landschulz, and S. L. McKnight.** 1988. In situ detection of sequence-specific DNA binding activity specified by a recombinant bacteriophage. Genes Dev. 2:807–806.
223. **Vojtek, A. B., S. M. Hollenberg, and J. A. Cooper.** 1993. Mammalian Ras interacts directly with the serine/threonine kinase Raf. Cell 74:205–214.
224. **Waksman, G., D. Kominos, S. C. Robertson, N. Pant, D. Baltimore, R. B. Birge, D. Cowburn, H. Hanafusa, B. J. Mayer, M. Overduin, M. D. Resh, C. B. Rios, L. Silverman, and J. Kuriyan.** 1992. Crystal structure of the phosphotyrosine recognition domain SH2 of V-src complexed with tyrosine-phosphorylated peptides. Nature (London) 358:646–653.
225. **Waksman, G., S. E. Shoelson, N. Pant, D. Cowburn, and J. Kuriyan.** 1993. Binding of a high affinity phosphotyrosyl peptide to the Src SH2 domain: crystal structures of the complexed and peptide-free forms. Cell 72:779–790.
226. **Waldron, C., and F. Lacroute.** 1975. Effect of growth rate on the amounts of ribosomal and transfer ribonucleic acids in yeast. J. Bacteriol. 122:855–865.
227. **Warne, P. H., P. R. Viciana, and J. Downward.** 1993. Direct interaction of Ras and the amino-terminal region of Raf-1 *in vitro*. Nature (London) 364:352–355.
228. **Weber, J., R. S.-F. Lee, S. Wilke-Mounts, E. Grell, and A. E. Senior.** 1993. Combined application of site-directed mutagenesis, 2-azido-ATP labeling, and *lin*-benzo-ATP binding to study the noncatalytic sites of *Escherichia coli* F_1-ATPase. J. Biol. Chem. 268:6241–6247.
229. **Weiel, J., and J. W. B. Hershey.** 1981. Fluorescence polarization studies of the interaction of *Escherichia coli* protein synthesis initiation factor 3 with 30S ribosomal subunits. Biochemistry 20:5859–5865.
230. **Weiel, J., and J. W. B. Hershey.** 1982. The binding of fluorescein-labeled protein synthesis initiation factor 2 to *Escherichia coli* 30 S ribosomal subunits determined by fluorescence polarization. J. Biol. Chem. 257:1215–1220.
231. **Wells, J. A., and H. B. Lowman.** 1992. Rapid evolution of peptide and protein binding properties *in vitro*. Curr. Opin. Struct. Biol. 2:355–362.
232. **Weng, Z., J. A. Taylor, C. E. Turner, J. S. Brugge, and C. Seidel-Dugan.** 1993. Detection of Src homology 3-binding proteins, including paxillin, in normal and v-Src-transformed Balb/c 3T3 cells. J. Biol. Chem. 268:14956–14963.
233. **Wensel, T. G., and L. Stryer.** 1986. Reciprocal control of retinal rod cyclic GMP phosphodiesterase by its γ subunit and transducin. Proteins Struct. Funct. Genet. 1:90–99.
234. **Wensel, T. G., and S. Stryer.** 1990. Activation mechanism of retinal rod cyclic GMP phosphodiesterase probed by fluorescein-labeled inhibitory subunit. Biochemistry 29:2155–2161.
235. **White, R. J., and S. P. Jackson.** 1992. The TATA-binding protein: a central role in transcription by RNA polymerases I, II and III. Trends Genet. 8:284–288.
236. **Whyte, P., K. J. Buchkovich, J. M. Horowitz, S. H. Friend, M. Raybuck, R. A. Weinberg, and E. Harlow.** 1988. Association between an oncogene and an anti-oncogene: the adenovirus E1A proteins bind to the retinoblastoma gene product. Nature (London) 334:124–129.
237. **Whyte, P., N. M. Williamson, and E. Harlow.** 1989. Cellular targets for transformation by the adenovirus E1A proteins. Cell 56:67–75.
238. **Wiedmann, M., D. Goerlich, E. Hartmann, T. V. Kurzchalia, and T. A. Rapoport.** 1989. Photocrosslinking demonstrates proximity of a 34 kDa membrane protein to different portions of preprolactin during translocation through the endoplasmic reticulum. FEBS Lett. 257:263–268.
239. **Wiedmann, M., T. V. Kurzchalia, E. Hartmann, and T. A. Rapoport.** 1987. A signal sequence receptor in the endoplasmic reticulum membrane. Nature (London) 328:830–833.
240. **Wiman, K. G.** 1993. The retinoblastoma gene: role in cell cycle control and cell differentiation. FASEB J. 7:841–845.
241. **Yanofsky, C., and M. Rochmeler.** 1958. The exclusion of free indole as an intermediate in the biosynthesis of tryptophan in *Neurospora crassa*. Biochim. Biophys. Acta 28:640–645.
242. **Yee, S.-P., and P. E. Branton.** 1985. Detection of cellular proteins associated with human adenovirus type 5 early region 1A polypeptides. Virology 147:142–153.
243. **Yong, H., G. A. Thomas, and W. L. Peticolas.** 1993. Metabolite-modulated complex formation between α-glycerophosphate dehydrogenase and lactate dehydrogenase. Biochemistry 32:11124–11131.
244. **Young, R., and R. Davis.** 1983. Yeast RNA polymerase II genes: Isolation with antibody probes. Science 222:778–782.
245. **Yu, H., M. K. Rosen, T. B. Shin, C. Seidel-Dugan, J. S. Brugge, and S. L. Schreiber.** 1992. Solution structure of the SH3 domain of Src and identification of its ligand-binding site. Science 258:1655–1668.
246. **Zebedee, S. L., C. F. Barbas III, Y.-L. Hom, R. H. Caothien, R. Graff, J. DeGraw, J. Pyati, R. LaPolla, D. R. Burton, R. A. Lerner, and G. B. Thornton.** 1992. Human combinatorial antibody libraries to hepatitis B surface antigens. Proc. Natl. Acad. Sci. USA 89:3175–3179.
247. **Zervos, A. S., J. Gyuris, and R. Brent.** 1993. Mxi1, a protein that specifically

interacts with Max to bind Myc-Max recognition sites. Cell 72:223–232.
248. **Zhang, X.-F., J. Settleman, J. M. Kyriakis, E. Takeuchi-Suzuki, S. J. Elledge, M. S. Marshall, J. T. Bruder, U. R. Rapp, and J. Avruch.** 1993. Normal and oncogenic p21^{ras} proteins bind to the amino-terminal regulatory domain of c-Raf-1. Nature (London) 364:308–313.
249. **Zhou, M., S. Felder, M. Rubinstein, D. R. Hurwitz, A. Ullrich, I. Lax, and J. Schlessinger.** 1993. Real-time measurements of kinetics of EGF binding to soluble EGF receptor monomers and dimers support the dimerization model for receptor activation. Biochemistry 32:8193–8198.
250. **Zimmerman, J. K., and G. K. Ackers.** 1971. Molecular sieve studies of interacting protein systems: behavior of small zone profiles for reversibly self-associating solutes. J. Biol. Chem. 246:7289–7292.

PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 5 :
C12Q 1/68, C12P 21/02
G01N 33/53, 33/531, A61K 35/14

A1

(11) International Publication Number: WO 91/04339

(43) International Publication Date: 4 April 1991 (04.04.91)

(21) International Application Number: PCT/US90/05155

(22) International Filing Date: 12 September 1990 (12.09.90)

(30) Priority data:
408,767 18 September 1989 (18.09.89) US

(71) Applicant: CALIFORNIA BIOTECHNOLOGY INC. [US/US]; 2450 Bayshore Parkway, Mountain View, CA 94043 (US).

(72) Inventors: CORDELL, Barbara ; 4051 Ben Lomond Drive, Palo Alto, CA 94306 (US). WOLF, David ; 2142 Bellview Drive, Palo Alto, CA 94303 (US).

(74) Agents: MURASHIGE, Kate, H. et al.; Irell & Manella, 545 Middlefield Road, Suite 200, Menlo Park, CA 94025 (US).

(81) Designated States: AT (European patent), AU, BE (European patent), CA, CH (European patent), DE (European patent)*, DK (European patent), ES (European patent), FR (European patent), GB (European patent), IT (European patent), JP, LU (European patent), NL (European patent), SE (European patent).

Published
With international search report.
Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.

(54) Title: ASSAYS AND REAGENTS FOR AMYLOID DEPOSITION

(57) Abstract

The present invention provides an in vitro tissue culture-based assay for amyloid deposition specific for Alzheimer's disease which is suitable for routine drug screening analysis. Immunological diagnostic reagents for Alzheimer's disease are also provided.

* See back of page

DESIGNATIONS OF "DE"

Until further notice, any designation of "DE" in any international application whose international filing date is prior to October 3, 1990, shall have effect in the territory of the Federal Republic of Germany with the exception of the territory of the former German Democratic Republic.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AT	Austria	ES	Spain	MC	Monaco
AU	Australia	FI	Finland	MG	Madagascar
BB	Barbados	FR	France	ML	Mali
BE	Belgium	GA	Gabon	MR	Mauritania
BF	Burkina Faso	GB	United Kingdom	MW	Malawi
BG	Bulgaria	GR	Greece	NL	Netherlands
BJ	Benin	HU	Hungary	NO	Norway
BR	Brazil	IT	Italy	PL	Poland
CA	Canada	JP	Japan	RO	Romania
CF	Central African Republic	KP	Democratic People's Republic of Korea	SD	Sudan
CG	Congo			SE	Sweden
CH	Switzerland	KR	Republic of Korea	SN	Senegal
CM	Cameroon	LI	Liechtenstein	SU	Soviet Union
DE	Germany	LK	Sri Lanka	TD	Chad
DK	Denmark	LU	Luxembourg	TG	Togo
				US	United States of America

ASSAYS AND REAGENTS FOR AMYLOID DEPOSITION

Field of the Invention

The present invention relates to assays and reagents useful for the chemical intervention of amyloidosis in Alzheimer's disease.

Background of the Invention

Alzheimer's disease (AD) is an age-related brain degenerative disease that is the most common cause of intellectual failure in late life. Neuritic or senile plaques and neurofibrillary tangles (NFT) are the hallmark characteristic of the histopathology of Alzheimer's brains. These plaques and tangles are believed to result from deposits of two different proteins which share the properties of the amyloid class of proteins specific for AD.

The major protein component of amyloid is an ~4 kilodalton (kd) protein, designated the beta-protein or A4 protein due to a partial beta pleated structure or its molecular weight, respectively. The amino acid sequence of A4 has been defined (Wong et al., (1985) Proc Natl Acad Sci USA 82:8729-8732) and full-length cDNA encoding a primary translation product of 695 residues has been cloned (Kang et al., (1987) Nature 325:733-736) while other cDNAs have been identified which encode a 751-residue or 770-residue precursor form (Ponte et al., (1988) Nature 331:525-527; Tanzi et al., (1988) Nature

331:528-530; and Kitaguchi et al., (1988) Nature 331:530-532).

The A4 protein accumulates extracellularly, both in brain parenchyma and in the walls of blood vessels, generally as amyloid plaques which form aggregate fibril structures and are insoluble on SDS-polyacrylamide gels. The fibrils are generally identified as amyloid based on their green birefringence after staining with Congo red and their 40- to 90-A diameter.

The second protein, mentioned previously, accumulates intracellularly in neurons of Alzheimer's brains (Castano and Frangione, (1988) Lab Invest 58:122-132) and forms tangles composed of structures resembling paired helical filaments (PHFs). In contrast to the beta-amyloid protein, the primary structure and number of proteins comprising PHFs are unknown. PHF-containing neurites are found in the periphery of the plaque, whereas deposits of beta-amyloid protein form the central core of mature plaques, surrounded by degenerated neurites and glial cells.

Although the etiology of AD is unknown, it has been demonstrated that the frequency of neuritic plaques found in the cortex of AD patients correlates with the degree of dementia (Roth et al., (1966) Nature 209:109-110; Wilcock and Esiri, (1982) J Neurol Sci 56:343-356). The therapeutic goals in amyloidosis are to prevent further deposition of amyloid material and to promote or accelerate its resorption. To date, there are no means available to treat the pathogenesis of AD and the paucity of understanding concerning the mechanism of amyloid formation in AD is a major obstacle in the development and design of therapeutic agents that can intervene in this process. Moreover, no animal models for brain amyloidosis with beta-amyloid protein deposits or PHFs exist, creating yet another obstacle to test such putative therapeutic agents.

Logical therapeutic approaches are now, however, emerging for treating the particular amyloidosis associated with AD. These approaches are attributable, in part, from the successes and failure gained in attempting to treat other forms of amyloidosis, such as the use of dimethyl sulfoxide which blocks amyloid formation from Bence Jones proteins in vitro (Coria et al., (1988) Lab Invest 58:454-458) and use of colchicine to reduce the size of renal amyloid deposits and induce clinical remissions in several cases of familial Mediterranean fever and amyloid nephropathy (Ravid et al., (1977) Ann Intern Med 87:568-570).

Efforts directed to the design of in vitro models of age-related cerebral amyloidogenesis using A4-derived synthetic peptides are disclosed in Castano et al., (1986) Biochem Biophys Res Comm 141:782-789, and in Kirschner et al., (1987) Proc Natl Acad Sci USA 84:6953-6957. Castano et al. demonstrated that amyloid fibrils could be formed in vitro when using a synthetic peptide corresponding to the amino-terminal 28 residues of the amyloid core protein. This 28 residue peptide, as well as a 17 residue sequence contained within the 28 amino acids, both formed fibrils which stain similarly to material isolated from Alzheimer's brains; however, the synthetic amyloid fibrils were soluble, unlike the naturally occurring insoluble amyloid isolated from Alzheimer's brains. Kirschner et al. demonstrated that the same 28 residue peptide could be produced as an insoluble aggregate; however, this particular in vitro model is not expected to correlate well to the cellular environment in which amyloid deposition occurs.

Dyrks et al., (1988) EMBO J 7:949-957 showed that a shortened cell-free translation product comprising the amyloid A42 part and the cytoplasmic domain of the 695-residue precursor can form multimers. While aggregation was observed employing an in vitro cell-free

system, this system fails to reveal whether aggregation of the translation product would naturally follow in vivo. Moreover, the in vitro cell-free system does not address protein stability issues, that is, whether adequate levels of the protein could be expressed, whether protein proteolysis exists, and other concerns generally associated with in vivo expression of recombinant proteins.

Therefore, there exists a need for a definitive cellular deposition model with which one may assay agents capable of chemically intervening in the process of amyloid deposition. Such a method should be relatively simple to perform and should be highly specific in distinguishing AD plaques from the plaques associated with other disorders. Furthermore, it is desirable that the assay be capable of being reduced to a standardized format. The present invention satisfies such needs and provides further advantages.

Summary of the Invention

The present invention provides a method for determining the ability of a potential therapeutic agent to intervene in the amyloid deposition process associated with Alzheimer's disease in a cellular environment, which method utilizes a recombinantly produced amyloid substrate in a screening assay. The present invention also allows for the development and use of immunological reagents to detect the formation of preamyloid protein aggregation in the cell lines provided by the invention.

To achieve the objects and in accordance with the purpose of the invention, as embodied and broadly described herein, a method of screening agents capable of intervention in Alzheimer's disease amyloidosis comprises:

a) culturing a cell line capable of expressing a gene encoding beta-amyloid protein under conditions suit-

able to produce the beta-amyloid protein as an insoluble, preamyloid aggregate;

b) combining a known quantity of the agent to be tested to the cell culture; and

c) monitoring the combination to determine whether preamyloid aggregate formation is reduced.

In an alternative embodiment of the invention, preamyloid formation can be induced through infection of a cell line with a recombinant virus capable of expressing the beta-amyloid protein as an insoluble preamyloid aggregate. Such recombinant viruses carry expression vectors comprising DNA encoding the beta-amyloid protein.

Immunoassay kits employing the reagents useful to screen potential amyloid intervening agents are also provided by the present invention.

Brief Description of the Drawings

FIG. 1 is a schematic illustration of two amyloid expression constructs employing the vaccinia pUV1 insertion vector.

FIG 2. illustrates the results of immunoprecipitation of ^{35}S-methionine labeled VV:A99 infected CV-1 cell lysates using APCP antibodies. The arrows mark A99 protein.

FIG 3. are fluorescent photomicrographs of infected CV-1 cells stained with APCP antibodies. FIG. 3A is a Mock control; FIG. 3B is a VV:CONT control; FIG. 3C is the VV:99 construct; and FIG. 3D is the VV:42 construct. The magnification is 200x with a 0.4 second exposure time for each photo.

FIG. 4 is a illustration of the modified beta-actin expression selection vector, pAX-neo, that was employed to express the beta-amyloid core constructs in mammalian cells.

Detailed Description of the Preferred Embodiments

As indicated above, the invention involves a method of screening agents capable of intervention in Alzheimer's disease amyloidosis.

As used herein, the term "beta-amyloid core protein" or "A4 protein" refers to an approximately 4 kd protein first identified by Glenner and Wong, (1984) Biochem Biophys Res Comm 120:885, which is defined at the amino terminus by sequence analysis as a mixture of four peptides with slightly different amino termini, the amino termini of the three smaller peptides being completely encoded by that of the largest.

The term "beta-amyloid precursor protein" refers to either the amyloid precursor protein of 695 amino acids (Kang et al., (1987) supra) or the 751 amino acid protein (Ponte et al., (1988) supra) containing within their sequence, the beta-amyloid core protein sequence defined above. The A4 core protein begins at amino acid 597 of the 695 amino acid protein and at amino acid 653 of the 751 amino acid sequence.

The terms "preamyloid aggregation", "preamyloid formation", or "preamyloid deposits" refer to a morphological description -- first discovered by Tagliavini et al., (1988) Neurosci Lett 93:191-196 -- of spherical, granular deposits which are considerably smaller than pre-plaques and plaques found at a high frequency in the brains of Alzheimer's victims. These deposits can be occasionally detected with silver stain but not with Congo red, a stain to which amyloid proteins demonstrate high binding affinity.

As used herein, the term "insertion vector" includes plasmids, cosmids or phages capable of mediating homologous recombination into a viral genome such that the DNA encoding the beta-amyloid protein is stably carried by the resulting recombinant virus. In one embodiment of the

invention plasmids constructed from vaccinia virus DNA are employed.

The term "expression vector" includes plasmids, cosmids or phages capable of synthesizing a protein encoded by the respective recombinant gene carried by said vector. Such vectors are independently replicated in or capable of integration into the chromosome of an appropriate host cell for expression of the amyloid protein.

A cell has been "transformed" by exogenous or heterologous DNA when such DNA has been introduced inside the cell. The transforming DNA may or may not be integrated (covalently linked) into chromosomal DNA making up the genome of the cell. In prokaryotes, yeast, and mammalian cells, for example, the transforming DNA may be maintained on an episomal element such as a plasmid. The cell has been stably transformed when the cell is able to establish cell lines or clones comprised of a population of daughter cells containing the transforming DNA. A "clone" is a population of cells derived from a single cell or common ancestor by mitosis. A "cell line" is a clone of a cell that is capable of stable growth in vitro for many generations.

A. Beta-Amyloid Coding Sequences

The beta-amyloid genes may be synthetic or natural, or combinations thereof. The gene encoding the natural 751 amino acid precursor protein is described in PCT WO88/03951, published 2 June 1988 and assigned to the same assignee of the present application, and the expression of the protein in mammalian cells is provided in Example 4 therein. The relevant portions of this publication are specifically incorporated herein by reference.

The genes encode the A42 core protein or an amyloid protein, A99, which comprises the A42 core protein and the cytoplasmic domain. This latter protein consists of the 42 residue core protein and 57 amino acids of the

cytoplasmic domain of the 751 precursor protein. The sequence of A99 is as follows:

```
                                        10
Asp Ala Glu Phe Arg His Asp Ser Gly Tyr Glu Val His His Gln
                20                                      30
Lys Leu Val Phe Phe Ala Glu Asp Val Gly Ser Asn Lys Gly Ala
                                        40      (42)
Ile Ile Gly Leu Met Val Gly Gly Val Val Ile Ala Thr Val Ile
                50                                      60
Val Ile Thr Leu Val Met Leu Lys Lys Lys Gln Tyr Thr Ser Ile
                                        70
His His Gly Val Val Glu Val Asp Ala Ala Val Thr Pro Glu Glu
                80                                      90
Arg His Leu Ser Lys Met Gln Gln Asn Gly Tyr Glu Asn Pro Thr
                                    (99)
Tyr Lys Phe Phe Glu Gln Met Gln Asn.
```

These genes are provided for expression of the desired protein using recombinant DNA expression vectors.

As mentioned above, these genes may be natural, synthetic or combinations thereof. When preparing a synthetic nucleotide sequence, it may be desirable to modify the natural amyloid nucleic acid sequence. For example, it will often be preferred to use codons which are preferentially recognized by the desired host. In some instances, it may be desirable to further alter the nucleotide sequence, either synthetic or natural, to create or remove restriction sites to, for example, enhance insertion of the gene sequence into convenient expression vectors or to substitute one or more amino acids in the resulting polypeptide to increase stability. A general method for site-specific mutagenesis is described in Noren et al., (1989) Science 244:182-188.

Peptides of this precursor protein, for example, those derived from the A4 core protein, are also provided herein for the generation of specific immunological re-

agents and may also be synthetic or natural. Synthetic oligonucleotides are prepared by either the phosphotriester method as described by Edge et al., (1981) Nature 292:756 and Duckworth et al., (1981) Nuc Acids Res 9:1691 or the phosphoramidite method as described by Beaucage and Caruthers, (1981) Tet Lett 22:1859 and Matteucci and Caruthers, (1981) J Am Chem Soc 103:3185, and can be prepared using commercially available automated oligonucleotide synthesizers.

B. Vaccinia Viral Vectors

The coding sequences for the amyloid proteins can be inserted into vaccinia virus plasmid insertion vectors for the purpose of generating recombinant vaccinia viruses using the methods described in Moss et al., (1983) Methods in Gene Amplification, Vol. 3, Elsevier-North Holland, p. 202-213; and in Moss et al., (1984) J Virol 49:857-864. The amyloid-vaccinia recombinants can then be used for (1) expression of the respective amyloid protein and analysis of preamyloid formation, and (2) production of amyloid antibodies.

The two vaccinia virus insertion vectors, pSC11 (Chakrabarti et al., (1985) Mol Cell Biol 5:3403-3409 and pUV1 (Falkner et al., (1987) Nuc Acids Res 15:7192) were used for the expression of the amyloid proteins and generation of amyloid-vaccinia recombinants. Both vectors are of the co-insertion variety and each contains two vaccinia virus promoters. One promoter (P1) is used to drive the expression of a selectable marker gene (in this case, beta-galactosidase) while the other promoter (P2) is used to drive expression of the heterologous amyloid DNA insert. Both are flanked by vaccinia virus DNA (an interrupted thymidine kinase [tk] gene) which facilitates homologous recombination into a wild-type vaccinia virus genome and provides a selection mechanism (generation of tk minus viruses). The pSC11 vector utilizes a vaccinia

early-late promoter (P7.5) to drive heterologous gene expression and has a single SmaI cloning site. The pUV1 vector utilizes a vaccinia late promoter (P11) to drive heterologous gene expression and is designed for the expression of fusion proteins behind the ATG of the P11 late gene. In all cases, amyloid-pUV1 constructs were made using the most 5' (after the ATG) cloning site (EcoRI) in order to avoid introduction of additional amino terminal amino acids into the native amyloid protein sequence.

C. Recombinant Expression Vectors and Hosts

It will also be understood by those skilled in the art that both procaryotic and eucaryotic systems may be used to express the amyloid genes described herein. Procaryotes most frequently are represented by various strains of E. coli; however, other microbial strains may also be used. Plasmid vectors which contain replication sites, selectable markers and control sequences derived from a species compatible with the host are used; for example, E. coli is typically transformed using derivatives of pBR322, a plasmid derived from an E. coli species by Bolivar et al., (1977) Gene 2:95. pBR322 contains genes for ampicillin and tetracycline resistance, and thus provides multiple selectable markers which can be either retained or destroyed in constructing the desired vector. Commonly used procaryotic control sequences which are defined herein to include promoters for transcription initiation, optionally with an operator, along with ribosome binding site sequences, include such commonly used promoters as the beta-lactamase (penicillinase) and lactose (lac) promoter systems (Chang et al., (1977) Nature 198:1056), the tryptophan (trp) promoter system (Goeddel et al., (1980) Nucleic Acids Res 8:4057), the lambda-derived P_L promoter (Shimatake et al., (1981) Nature 292:128) and N-gene ribosome binding site, and the

trp-lac (trc) promoter system (Amann and Brosius, (1985) Gene 40:183).

In addition to bacteria, eucaryotic microbes, such as yeast, may also be used as hosts. Laboratory strains of Saccharomyces cerevisiae, Baker's yeast, are most used although a number of other strains or species are commonly available. Vectors employing, for example, the 2 micron origin of replication of Broach, (1983) Meth Enz 101:307, or other yeast compatible origins of replication (see, for example, Stinchcomb et al., (1979) Nature 282:39; Tschumper et al., (1980) Gene 10:157 and Clarke et al., (1983) Meth Enz 101:300) may be used. Control sequences for yeast vectors include promoters for the synthesis of glycolytic enzymes (Hess et al., (1968) J Adv Enzyme Reg 7:149; Holland et al., (1978) Biochemistry 17:4900). Additional promoters known in the art include the promoter for 3-phosphoglycerate kinase (Hitzeman et al., (1980) J Biol Chem 255:2073). Other promoters, which have the additional advantage of transcription controlled by growth conditions and/or genetic background are the promoter regions for alcohol dehydrogenase 2, isocytochrome C, acid phosphatase, degradative enzymes associated with nitrogen metabolism, the alpha factor system and enzymes responsible for maltose and galactose utilization. It is also believed terminator sequences are desirable at the 3' end of the coding sequences. Such terminators are found in the 3' untranslated region following the coding sequences in yeast-derived genes.

It is also, of course, possible to express genes encoding polypeptides in eucaryotic host cell cultures derived from multicellular organisms. See, for example, Axel et al., U.S. Patent No. 4,399,216. These systems have the additional advantage of the ability to splice out introns and thus can be used directly to express genomic fragments. Useful host cell lines include VERO, HeLa, baby hamster kidney (BHK), CV-1, COS, MDCK, NIH 3T3, L,

and Chinese hamster ovary (CHO) cells. Expression vectors for such cells ordinarily include promoters and control sequences compatible with mammalian cells such as, for example, the commonly used early and late promoters from Simian Virus 40 (SV40) (Fiers et al., (1978) Nature 273:113), or other viral promoters such as those derived from polyoma, herpes virus, Adenovirus 2, bovine papilloma virus, or avian sarcoma viruses. The controllable promoter, hMTII (Karin et al., (1987) Nature 299:797-802) may also be used. General aspects of mammalian cell host system transformations have been described by Axel, supra.

Insect expression systems may also be employed to express the amyloid genes. For example, the baculovirus polyhedrin gene has been employed for high-level expression of heterologous proteins (Smith et al., (1983) Mol Cell Biol 3(12):2156-2165; Summers et al., "Genetic Engineering of the Genome of the Autographa Californica Nuclear Polyhedrosis Virus", Banbury Report: Genetically Altered Viruses in the Environment, 22:319-339, Cold Spring Harbor Laboratory, 1985).

D. Generation of Stably Transfected Cell Lines

The amyloid DNA clones expressed in vaccinia can also be used to generate stably transfected cell lines expressing the amyloid proteins. In general, these cell lines are generated by first constructing one of two expression plasmids. In both expression plasmids, the selectable marker is provided by a G418 neomycin expression cassette (neo) consisting of the SV40 early promoter, the bacterial kanamycin-resistance gene also containing its own promoter, the SV40 intervening sequence, and the SV40 polyadenylation site from the early region. In the first expression plasmid, the amyloid DNA cloning site is flanked at the 5' end by the human metallothionein gene promoter, pMtIIa, modified with an SV40 enhancer, and at the 3' end by the SV40 polyadenylation site from the early

region. In the second expression construct, the amyloid DNA cloning site is flanked at the 5' end by a beta-actin promoter, and at the 3' end by a sequence encoding a useful polyadenylation site, such as that of the SV40 early region or the beta-actin gene.

Each of the vectors described above can be transformed into a mammalian cell line such as, but not limited to, those described in the following examples by either calcium phosphate-DNA coprecipitation or electroporation. A day later, the cells are subjected to 1 mg/ml G418 to provide pools of G418-resistant colonies. Successful transformants, also having a stable inheritance of the DNA contained in the expression construct, are then plated at low density for purification of clonal isolates. Clonal isolates are then analyzed for maximum production of the amyloid protein of interest and high-producing clones are expanded to serve as stock.

E. Detection Methods for Preamyloid Formation

The diagnosis of amyloidosis is established by demonstration of the characteristic emerald-green birefringence of tissue specimens stained with Congo red and examined by polarization microscopy. Congo red staining is generally carried out using commercially available diagnostic kits. The isolation and characterization of the A4 protein has allowed specific antibodies to be raised that recognized cerebral amyloid in Alzheimer's disease (Allsop et al (1986) Neurosci Lett 68:252-256). Moreover, Tagliavini et al., (1988) supra, have demonstrated that antibodies can be generated which detect in both Alzheimer's patients and to a lesser extent in non-demented individuals preamyloid deposits, which deposits lack the tinctorial and optical properties of amyloid and are, therefore, undetectable using conventional staining methods employing principally Congo red, but also thioflavin S or silver salts.

Standard protocols can be employed for preparing antibodies directed against the amyloid proteins of the invention. Techniques for preparing both polyclonal and monoclonal antibodies are well known in the art. Briefly, polyclonal antibodies are prepared by injecting amyloid protein or synthetic amyloid peptides with an adjuvant into an animal such as rabbits or mice. The amyloid protein may need to be conjugated to a carrier protein such as bovine serum albumin or keyhole limpet hemacyanin using a chemical process which employs carbodiimide, glutaraldehyde, or other cross-linking agents. Alternatively, the protein may be administered without being conjugated to a carrier protein. Vaccinia virus recombinants which are expressing amyloid proteins may also be used to prepare antibodies. The vaccinia virus recombinants are injected into an animal and then the animal is boosted several weeks after the initial immunization. Ten days to two weeks later the animals are bled and antiserum is collected and analyzed for titer.

Monoclonal antibodies are commonly prepared by fusing, under appropriate conditions, B-lymphocytes of an animal which is making polyclonal antibodies with an immortalizing myeloma cell line. The B-lymphocytes can be spleen cells or peripheral blood lymphocytes. Techniques for fusion are also well known in the art, and in general, involve mixing the cells with a fusing agent such as polyethylene glycol. Successful hybridoma formation is assessed and selected by standard procedures such as, for example, HAT medium. From among successful hybridomas, those secreting the desired antibody are screened by assaying the culture medium for their presence.

Standard immunological techniques such as ELISA (enzyme-linked immunoassay), RIA (radioimmunoassay), IFA (immunofluorescence assay) and Western blot analysis, which are well known in the art, can be employed for diagnostic screening for amyloid expression. A vast

literature now exists with respect to various modifications of the basic assay principle, which is simply that there must be a specific association between target analyte and antibody, which association is detectable qualitatively and/or quantitatively. Fluorescent, enzymatic, or radioactive labels are generally used.

One typical arrangement utilizes competition, between labeled antigen (e.g. amyloid protein) and the analyte, for the antibody, followed by physical separation of bound and unbound fractions. Analyte competes for the binding of the labeled antigen; hence more label will remain in the unbound fraction when larger amounts of analyte are present. In this competitive-binding type assay, the sample is incubated with a known titer of labeled amyloid protein and amyloid protein antibody. Antibody-protein complex is then separated from uncomplexed reagents using known techniques and the amount of label in the complexed material is measured, e.g. by gamma counting in the case of radioimmunoassay or photometrically in the case of enzyme immunoassay. The amount of amyloid protein in the sample, if any, is determined by comparing the measured amount of label with a standard curve.

Other embodiments of this basic principle include use of labeled antibodies per se, sandwich assays involving a three-way complex between analyte, anti-analyte antibody, and anti-antibody wherein one of the components contains a label, and separation of bound and unbound fractions using an immunosorbent. Agglutination assays which result in visible precipitates are also available (Limet et al., (1982) J Clin Chem Clin Biochem 20:142-147).

F. Screening Assay

The present assay provides one of the first steps in addressing the question whether preamyloid corti-

cal deposits herald organic dementia. The concomitant appearance of preamyloid deposits and senile plaques suggests that preamyloid deposits may evolve into senile plaques.

Down's syndrome is the one known disease closely related to the proposed etiology of AD. As from their twenties onward, Down's patients develop the full spectrum of Alzheimer's changes, i.e., NFTs, congophilic angiopathy and senile plaques. As reported in Giacione et al., (1989) Neurosci Letts 97:232-238, a time-related analysis of preamyloid deposits and senile plaque distribution showed an age-dependent, inverse correlation between extracellular preamyloid deposits and senile plaque in Down's patients. While a similar, time-dependent study with Alzheimer's patients cannot be conducted, it is expected that a corresponding pattern (preamyloid turning to senile plaque deposits) would be found. Therapeutic agents which interfere with this process promise the development of successful therapeutic regimens for Alzheimer's disease.

In the practice of the method of the invention, the expression of the amyloid protein is initiated by culturing the transformed cell line under conditions which are suitable for cell growth and expression of the amyloid protein. In this method, high level expression of the protein is preferred. In one embodiment of the invention, a CHO cell line transformed with a beta-actin vector comprising the DNA encoding the A42 or A99 amyloid protein is grown in a mammalian culture medium such as, for example, a 1:1 mixture of F12 medium and DME medium with 10% fetal calf serum for 5-72 hr at 37°C. Transfected viral monolayers are selected and plaque purified, and stocks of amyloid-vaccinia recombinant viruses are prepared.

The formation of the preamyloid aggregates can be monitored by standard immunocytochemical methods using,

for example, beta-amyloid primary antibodies which are detected using a secondary, labeled anti-antibody. If one is interested in testing whether the compound of interest can inhibit preamyloid formation, the compound is introduced to the culture medium before monitoring for preamyloid aggregation. Alternatively, the compound is introduced to the culture medium after preamyloid formation has been established and this reaction mixture is monitored to see whether the compound induces amyloid resorption.

Potential therapeutic compounds for use in the present invention include, for example, amyloid-fibril denaturing agents such as dimethyl sulfoxide, and cytotoxic agents such as colchicine and chlorambucil. The efficacy of these agents may be monitored through observation of reduced antibody binding to the amyloid deposit. Reduction in such binding is indicative of reduced preamyloid deposition. Alternatively, preamyloid formation in the host cell may trigger other cellular events which could be employed as markers unrelated to the etiology of Alzheimer's disease, but correlative with the presence of preamyloid deposits. For example, an increase in the level of certain enzymes, specifically proteases, may be measured in lieu of the preamyloid deposition. Typically, an increase in the concentration levels of these enzymes is observed when cultured cells are subjected to stress.

The present invention also encompasses kits suitable for the above diagnostic or screening methods. These kits contain the appropriate reagents and are constructed by packaging the appropriate materials, including the preamyloid protein aggregates immobilized on a solid support with labeled antibodies in suitable containers, along with any other reagents (e.g., wash solutions, enzyme substrate, anti-amyloid antibodies) or other materials required for the conduct of the assay.

The reagents are usually premeasured for ease of use. An optional component of the kit is a set of instructions describing any of the available immunoassay methods. For example, a kit for a direct assay can comprise preamyloid proteins aggregates immobilized on a solid immunoassay support and a container comprising labeled antibody to the amyloid protein, as well as the other reagents mentioned above.

The following examples are designed to elucidate the teachings of the present invention, and in no way limit the scope of the invention. Most of the techniques which are used to transform cells, construct vectors perform immunoassays, and the like are widely practiced in the art, and most practitioners are familiar with the standard resource materials which describe specific conditions and procedures. The examples are written in observation of such knowledge and incorporate by reference procedures considered conventional in the art.

EXAMPLE 1

Description of Amyloid Plaque Core DNA Constructs

The following examples describe the expression vectors containing the 42 amino acid plaque core region (A42), and the 42 amino acid plaque core region including the 57 amino acid adjacent carboxy-terminal region of the beta-amyloid precursor protein (A99). Alternative constructs for the A42 and A99 constructs were prepared which included a 17 amino acid amyloid signal sequence. As these constructs did not express the amyloid protein well, further experimentation with these vectors was not performed.

Recombinant vaccinia viruses bearing amyloid DNAs encoding each of the two amyloid constructs (VV:A42 and VV:A99) were generated by standard methods as reviewed by Mackett and Smith in (1986) J Gen Virol 67:2067-2082,

which is incorporated herein by reference. FIG. 1 illustrates the various amyloid expression constructs, all of which were modified at the 5' end to satisfy the cloning constraints of the vaccinia P11 promoter in the pUV1 vector. Specifics for each construct are as follows:

A. VV:A42:

The A42-encoding sequence (nucleotides 2080 to 2205, numbered in accordance with the 751 amyloid precursor sequence) was synthesized as a 145 basepair (bp) EcoRI-BamHI oligomer, provided below, containing the appropriate TGA stop codon and an amino-terminal Asn-Ser adaptor sequence:

5' AAT TCC GAT GCA GAA TTC CGA CAT GAC TCA
GGA TAT GAA GTT CAT CAT CAA AAA TTG GTG
TTC TTT GCA GAA GAT GTG GGT TCA AAC AAA
GGT GCA ATC ATT GGA CTC ATC GTG GGC GGT
GTT GTC ATA GCG TGA TCT AGA TGA G 3'

The synthetic fragment was ligated to EcoRI- and BamHI-digested pGem1 (Promega-Biotec), deriving pGemA42. The EcoRI-BamHI fragment of pGemA42 was subsequently isolated and ligated into the EcoRI-BamHI site of pUV1 deriving pUV1:A42.

The XbaI-SalI fragment of pUV1:A42 (287bp) was further subcloned into mp18 for sequence confirmation.

B. VV:A99:

The DNA encoding the amyloid protein for the pUV1-A99 constructs was derived from 4T4B, a plasmid encoding the 751 amino acid precursor protein. The construction of plasmid 4T4B is described in Example 3 of PCT/US87/02953, owned by the same assignee. The relevant portions of this publication are incorporated herein by reference. The 590 bp DdeI-PvuII fragment of plasmid 4T4B

was isolated from the carboxy-terminal 1 kilobase (kb) EcoRI fragment of 4T4B and ligated with a 27 bp EcoRI-DdeI adaptor sequence and cloned into the EcoRI- and SmaI-digested pUV1, deriving pUV1:A99.

The 761 bp XbaI-SalI fragment of pUV1:A99 was further subcloned into the XbaI-SalI vector fragment of mp18 and pGem2. Sequence data confirmed the predicted sequence.

EXAMPLE 2
Expression of Amyloid Proteins

The vaccinia insertion vectors described in Example 1 were used to generate amyloid-vaccinia recombinant viruses as follows.

A. Preparation of Amyloid-Vaccinia Virus Recombinants

Confluent monolayers of CV-1 cells in 60 mm dishes were infected with vaccinia virus (Wyeth strain) at a multiplicity of infection (moi) of 0.05 pfu/cell. At 0.5 hr post-infection, the cells were transfected with a calcium phosphate precipitate of 10 ug insertion plasmid DNA and 0.5 ug wild-type vaccinia virus DNA. Cells were fed with complete medium and incubated at 37^{o}C for two days. Monolayers were collected and TK- vaccinia viruses were selected on TK-143 cells in the presence of 5-bromodeoxyuridine (BudR) at 25 ug/ml. At 48 hr after infection, monolayers were overlaid with 1% agarose containing 300 ug/ml 5-bromo-4-chloro-3-indolyl-B-D-galactopyranoside (Xgal). At 4-6 hr, blue plaques were picked and further purified by two additional rounds of plaque purification in the presence of BudR and Xgal. Stocks of the amyloid-vaccinia recombinant viruses were prepared in TK-142 or CV-1 cells. Recombinant viral DNA was prepared from each stock and was shown by Southern blot analysis to contain the appropriate amyloid DNA

insert and to be free of contamination with wild-type or spontaneous TK- vaccinia.

B. Identification of Amyloid-specific Polypeptides Produced By Vaccinia Virus Recombinants

Characterization of the CV-1 expressed VV:A42 and VV:A99 amyloid proteins was carried out employing immunoprecipitation and polyacrylamide gel analysis of ^{35}S-methionine-labeled infected cell protein using antibodies directed against the carboxy-terminal region of the amyloid precursor.

The beta-amyloid antibodies were generated from synthetic peptides. The synthetic peptides were prepared using solid phase synthesis according to standard protocols. Purification of the crude peptides was accomplished by desalting with gel filtration followed by ion-exchange chromatography and preparative reverse-phase liquid chromatography. Each peptide was fully characterized by amino acid composition and sequence analysis. COOH-CORE corresponds to amino acids 653-680(DAEFRHDSGYEVHHQKLVFFAEDVGSSA) (the carboxy-terminal two amino acids were taken from the amino acid sequence of Masters et al., (1985) Proc Natl Acad Sci 82:4245-4249 and are different in the deduced translation of the A4 cDNA of Ponte et al., supra. COOH-B2 and COOH-C2 correspond to amino acids 736-751(NGYENPTYKFFEQMQN), COOH-B3 and COOH-C3 correspond to amino acids 705-719(KKKQYTSIHHGVVEV) and COOH-C5 corresponds to amino acids 729-742(HLSKMQQNGYENPT). Reference for the numbering of peptides along the topology of the A4 precursor is from Ponte et al., supra. New Zealand white rabbits were immunized intradermally with 500 ug of peptide conjugated to keyhole limpet hemocyanin. The rabbits were first bled at 4 weeks and 1 week later the rabbits were boosted with 250 ug conjugated peptide. Subsequent bleeds were done at 3 week intervals with boosts following 1 week later. All

animals were treated in accordance with institutional guidelines. Antibody titers against the appropriate peptide were determined by enzyme-linked immunosorbent assays coupled with horseradish peroxidase and found to be $7.4x10^4$, $2.7x10^5$, $1x10^5$, $9.1x10^6$, $8.2x10^5$, and $2.5x10^5$ for COOH-CORE, COOH-B2, COOH-C2, COOH-B3, COOH-C3, and COOH-C5, respectively.

Antibodies to 9523 correspond to amino acids 673-685(AEDVGSKNGAIIG) and 9524 correspond to amino acids 701-712(LVMLKKKQYTSI). Antibodies to these two peptides were generated by coinjecting New Zealand white rabbits each with 200 ug methylated bovine serum albumin (PBS) plus 200 ug of the respective synthetic peptide in PBS. Rabbits were boosted one, two and three weeks following primary inoculation with identical amounts of peptide. Serum samples were taken at week 6 and titered against APCP synthetic peptide. Titers achieved were $1.5x10^4$ for 9523 and $4x10^5$ for 9524.

CV-1 cells were infected with VV:99 at a multiplicity of infection of one. ^{35}S-methionine (250 uCi/ml) was added at 20 hr post infection for 4 hr. Cell lysates were prepared and aliquots containing 10^7 cpm were immunoprecipitated with amyloid-specific antisera (COOH-B3, COOH-C5 and COOH-CORE) or normal rabbit serum and protein A.

Immunoprecipitates of ^{35}S-methionine cell lysates were analyzed on denaturing 20% SDS-polyacrylamide gels. As shown in FIG. 2, high levels of expression and stability of the A99 protein generated by VV:A99 was demonstrated. The control sera (normal, nonimmune rabbit sera) did not display reactivity with the VV:A99 protein product. The VV:A99 amyloid core protein migrated as a broad band spanning approximately 11.5-17 kd molecular weight. In addition, higher molecular weight forms of the A99 protein were clearly observed.

The expression product of VV:A99 demonstrated high level expression of the 99 amino acid core protein and showed evidence of self-aggregation as well as aggregation with other proteins or self-aggregation combined with proteolysis since multimers of A99 did not always occur in integers of 11.5-17 kd.

EXAMPLE 3

Staining of A42 and A99 Expressing Cells

Two human, SK-N-MC (ATCC # HTB10) and IMR-32, (ATCC # CCL127) and one rat, PC-12 (Green and Tischler, (1976) Proc Natl Acad Sci USA 73:2424-2428) neuronal cell lines were examined for their ability to permit efficient infection with the VV:A42 and VV:A99 recombinant viruses. All cell lines were documented as permissive hosts for vaccinia virus replication by infecting cells with a given amount (moi=2) of vaccinia virus of known titer. The infected cells were harvested 20 hours after infection, disrupted by freeze-thaw, and then titered. The yield was compared to the input viral units and if 20-100 fold increase results, the host cell was considered permissible for vaccinia replication.

These neuronal lines and the CV-1 cell line were employed for amyloid staining studies. The culture medium for each host was as follows:

CV-1: The medium was Eagle MEM supplemented with 10% FBS, penicillin, streptomycin and L-Gln.

SK-N-MC: Eagle MEM supplemented with 10% FBS, non-essential amino acids, penicillin, streptomycin and L-Gln.

PC-12: DMEM21, 5% DHS, 5% DFBS and L-Gln; and

IMR-32: Eagle MEM (Hank's BSS) and 10% deltaFBS plus nonessential amino acids, penicillin, streptomycin and L-Gln.

Each cell line was grown to confluency on a microscope slide divided into 4 individual chambers (Lab Tech). One chamber was mock infected, the second infected with a control recombinant virus lacking A4 sequences (VV:CONT), the third chamber infected with VV:A99, and the fourth chamber infected with VV:A42. This is an internally controlled method since each slide was manipulated as a single unit.

Viral infections were carried out at a moi from 5 to 20 viral plaque forming units (pfu) per cell and were harvested for staining at approximately 20 hours post infection. Slides prepared for immunocytochemistry were fixed with 4% paraformaldehyde and permeabilized with 0.2% Triton X-100 prior to treatment with primary and rhodamine-conjugated second antibodies (Capell Labs). Briefly, after permeabilization, cells were washed with PBS containing 0.2% gelatin. 100 ul of primary amyloid antibody (diluted 1/200 with PBS plus 0.2% gelatin) was incubated on the cells at 37^{o}C for 30 minutes. Cells were washed for 10 minutes in PBS and 0.2% gelatin, then incubated at 37^{o}C for 20 minutes with a 1/200 dilution (in PBS and gelatin) of secondary antibody (goat-anti-rabbit) tagged with Rhodamine. Cells were washed for 10 min in PBS and gelatin, then mounted for visualization in a fluorescent microscope. Antibodies used with success included 9523, 9524, B3 and C5. CORE antibodies were not assessed. Alternatively, the slides were fixed in 4% paraformaldehyde then stained with Thioflavin S or Congo red, and counterstained with hematoxylin according to directions in commercial kits (Sigma).

IMR-32 and PC-12 cells presented some technical difficulties and thus further investigation with these cell lines was terminated. The IMR-32 cells did not adhere well to the microscope slides, which could be alleviated by pretreatment with laminin, and, moreover, the IMR-32 cells did not tolerate the serum-free conditions during the infections. PC-12 cells showed high background immunostaining, hence, differences between experimental and control samples were not dramatic.

FIG. 3 shows fluorescent photomicrographs of CV-1 cells stained with 1/200 dilutions of the core domain antibodies 9523 antibodies. Specific and robust staining was seen in only the VV:A99 and VV:42 infected cells. VV:99 specific staining, but not VV:42 staining, was seen with the B3 antibody as would be anticipated since this region is not included in the VV:A42 construct (results not shown). Faint punctate staining was observed for both antibodies on all cells presumably due to endogenous A4 precursor expression. The VV:A99 and VV:A42 infected cells displayed strong reactivity in the form of large deposit-like structures which are cell associated. The deposit-like structures are probably not cell debris from the viral cytopathicity since they are not seen in the VV:CONT cells and their immunoreactivity could be eliminated by preadsorption of the antisera with the synthetic peptide used to raise the serum.

The possible potentiating effect of aluminum on deposit formation was investigated by pretreating the cells with 50 mM $AlCl_3$. Aluminum might be considered a "cofactor" in the pathology of amyloid formation since it is present in plaques. However, no obvious qualitative difference in the degree of deposit formation between cultures treated and untreated with aluminum was found.

It seems relevant that several researchers investigating A4 core domain immunoreactivity in brains of Alzheimer's victims describe similar structures as those

in FIG. 3. Each group reported finding significant amounts of specifically stained spherical, granular deposits which were considerably smaller than pre-plaques and plaques (Davies et al., (1988) Neurolog 38:1688-1693; Ikeda et al., (1989) Lab Invest 60:113-122; Tagliavini et al., (1988) supra; Tate-Ostroff et al., (1989) Proc Natl Acad Sci 86:745-749). All research groups independently propose that the observed small granular deposits are the very early stages of amyloid plaque development. The structures observed in our cell culture system are analogous to those seen in the Alzheimer's diseased brain. It was noted by these investigators that the granular deposits could be occasionally detected with silver stain but not with Congo red. Because the Alzheimer's granular deposits were highly reactive with A4 antisera but were not easily reacted with stains capable of recognizing the tinctorial properties of amyloid, the structures were termed "preamyloid" deposits.

EXAMPLE 4

Establishment of Stable Cell Lines

A number of constructs expressing the beta-amyloid core protein were constructed using a derivative of the beta-actin expression/selection vector designated pHbetaAPr-1-neo. This vector, illustrated in FIG. 4, is a combination of the following elements:

a) bp 1-4300 is the 4.3 kb EcoRI-AluI fragment from the human beta-actin gene isolate p14Tbeta-17 (Leavitt et al., (1984) Mol Cell Biol 4:1961-1969). For sequencing details of the promoter see Ng et al., (1985) Mol Cell Biol 5:2720-2732. The cap site, 5' untranslated region and IVS 1 positions are indicated in FIG. 4. There is no ATG codon present in the 5' UT nor in the polylinker region from the 3' splice site to the BamHI site;

b) bp 4300-4320 is in part derived from pSP64 polylinker (Melton et al., (1984) Nuc Acids Res 12:7035-7056);

c) bp 4320-6600 is derived from pcDV1 (Okayama & Berg, (1983) Mol Cell Biol 3:280-289); and

d) bp 6600-10000 is the PvuII-EcoRI fragment from pSV-neo (Southern & Berg, (1982) J Mol App Genet 1:327-341) containing the bacterial neomycin gene linked to the SV40 origin plus early promoter. The direction of transcription is as indicated in FIG. 4. This vector was altered by deleting the EcoRI site and adding a new EcoRI site within the polylinker 3' to the SalI site and 5' to the HindIII site. This modified vector is designated pAX-neo. Beta-actin A42 was constructed by excising the EcoRI-BamHI 145 bp fragment from pGEM-A42, adding a SalI-EcoRI adaptor sequence (5'-TCG ACA TGG ATG CAC AAT TA-3') and cloning into the pAX-neo expression vector at the SalI-and BamHI sites. The beta-actin A99 plasmid was constructed by excising the 670 bp EcoRI-HindIII fragment of $pGEM_2$-A99, adding the above-described SalI-EcoRI adaptor sequence and cloning into the pAX-neo vector at the SalI and HindIII sites.

Each construct was introduced into CHO cells by the calcium phosphate precipitation method using 7 ug of each DNA per 10^6 cells, and a resistant population was selected with G418-neomycin. The efficiency of transfection for the A99 or A42 constructs was over 10^3 for 10^6 cells and pools of cells transfected with either beta-actin A99 or with beta-actin A42 were selected using G418-neomycin resistance (500 ug/ml).

Cell lysates from these pools are prepared and analyzed by immunoprecipitation of the A4 proteins as well as by Western blotting. High expressing clones are then selected and assayed for "preamyloid" deposits using the immunocyto-staining procedures described in Example 3.

EXAMPLE 5

Assay for Preamyloid Deposition

Cells infected with VV:99 or VV:42 which are capable of forming amyloid deposits are plated in a 96-well microtiter plate. To make the appropriate dilutions and additions, an automated pipetter is used to introduce the drug to be tested to the cells. A range of concentrations of the drug is incubated in a tissue culture incubator (or preincubated) with the cells at 37°C for a predetermined time period, or alternatively, for 3 to 72 hours.

Following incubation, the culture media is removed, and the cells are prepared for preamyloid measurement as follows. The cells are fixed for immunocytochemical staining with amyloid antibodies. The primary antibodies are introduced followed by incubation with labeled, secondary anti-antibodies and the level of binding between the primary and secondary antibodies is measured using an ELISA plate reader to record the optical density of the labeled antibody. A smaller optical density reading as compared to a control sample of cells grown in the absence of the test drug is indicative of that drug's ability to inhibit amyloid deposition. This procedure may be modified to permit detection of preamyloid dissolution using a correlative enzyme marker.

It will be apparent to those skilled in the art that various modifications and variations can be made in the method of the present invention without departing from the scope or spirit of the invention. Thus, it is intended that the claims cover the modifications and variations of the invention.

What is claimed is:

1. A method of screening agents capable of intervention in Alzheimer's disease amyloidosis comprising:

a) culturing a cell line capable of expressing a gene encoding beta-amyloid protein under conditions suitable to produce the beta-amyloid protein as an insoluble, preamyloid aggregate;

b) combining a known quantity of the agent to be tested to the cell culture; and

c) monitoring the combination to determine whether preamyloid aggregate formation is reduced.

2. The method of claim 1 wherein the beta-amyloid gene encodes a protein comprising the amyloid plaque core domain.

3. The method of claim 1 wherein the beta-amyloid gene encodes a protein comprising the amyloid plaque core and the carboxy-terminal domains.

4. The method of claim 3 wherein the beta-amyloid gene encodes the following polypeptide:

10
Asp Ala Glu Phe Arg His Asp Ser Gly Tyr Glu Val His His Gln
20 30
Lys Leu Val Phe Phe Ala Glu Asp Val Gly Ser Asn Lys Gly Ala
40 (42)
Ile Ile Gly Leu Met Val Gly Gly Val Val Ile Ala Thr Val Ile
50 60
Val Ile Thr Leu Val Met Leu Lys Lys Lys Gln Tyr Thr Ser Ile
70
His His Gly Val Val Glu Val Asp Ala Ala Val Thr Pro Glu Glu
80 90
Arg His Leu Ser Lys Met Gln Gln Asn Gly Tyr Glu Asn Pro Thr
(99)
Tyr Lys Phe Phe Glu Gln Met Gln Asn.

5. The method of claim 1 wherein the preamyloid is derived from stable cell lines infected with recombinant vaccinia virus comprising the beta-amyloid gene.

6. The method of claim 1 wherein the cell lines are derived from mammalian host cells.

7. The method of claim 1 wherein the agent to be tested is introduced during the growth phase of the cell culture to determine whether the agent inhibits preamyloid plaque formation.

8. The method of claim 1 wherein the agent to be tested is added to the amyloid aggregate to determine whether the agent dissolves preamyloid plaque formation.

9. An immunological reagent capable of detecting preamyloid aggregate formation.

10. The immunological reagent of claim 9 which is a monoclonal antibody.

11. The immunological reagent of claim 9 which is a polyclonal antibody.

12. A kit for an immunoassay to screen compounds capable of chemical intervention in amyloidosis of Alzheimer's disease comprising:

a predetermined amount of preamyloid aggregate specific for Alzheimer's disease; and

a predetermined amount of labeled antibody to said preamyloid aggregate.

13. The kit according to claim 12 wherein said label is a component of an enzymatic reaction.

14. The kit according to claim 12 wherein said preamyloid aggregate is immobilized on a solid immunoassay support.

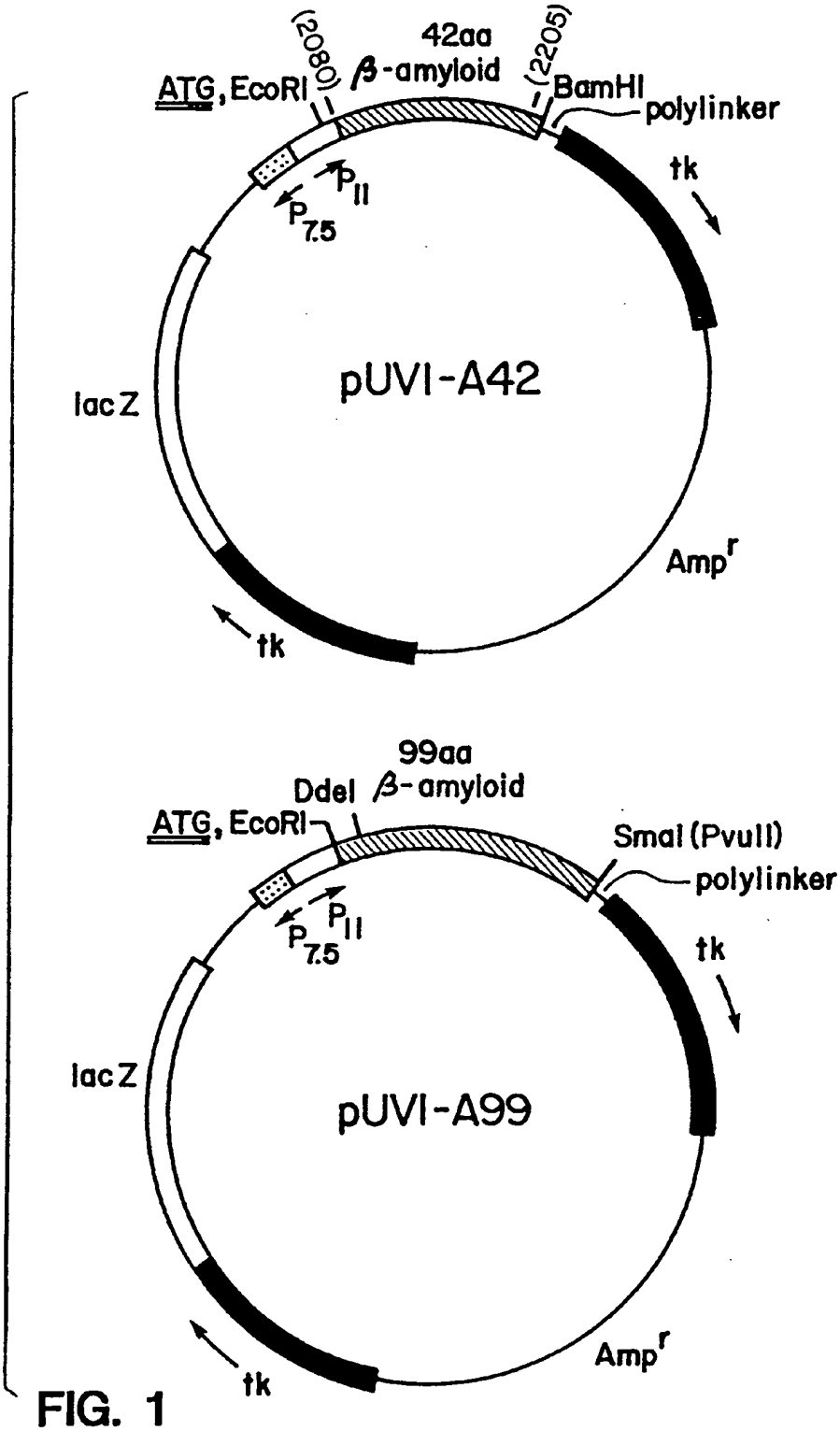

(2080)
42aa
β-amyloid
(2205)
ATG, EcoRI
BamHI
polylinker
tk
P11
P7.5
lacZ
pUVI-A42
Amp^r
tk
99aa
DdeI
β-amyloid
ATG, EcoRI
SmaI (PvuII)
polylinker
tk
P11
P7.5
lacZ
pUVI-A99
Amp^r
tk

FIG. 1

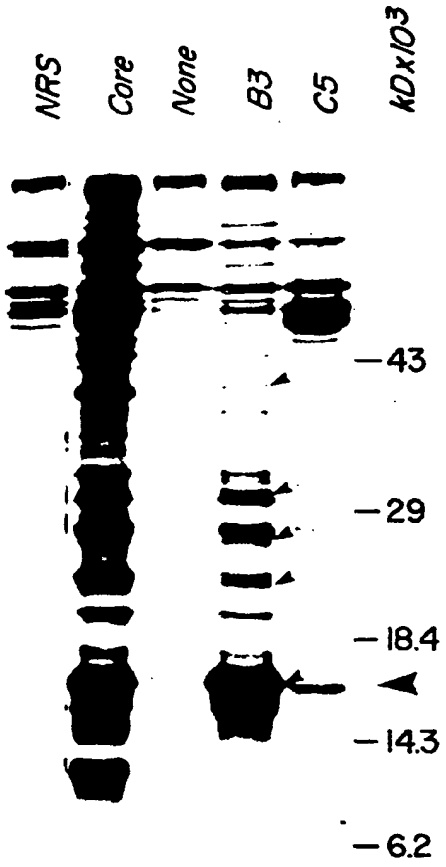


FIG.2

FIG.3A

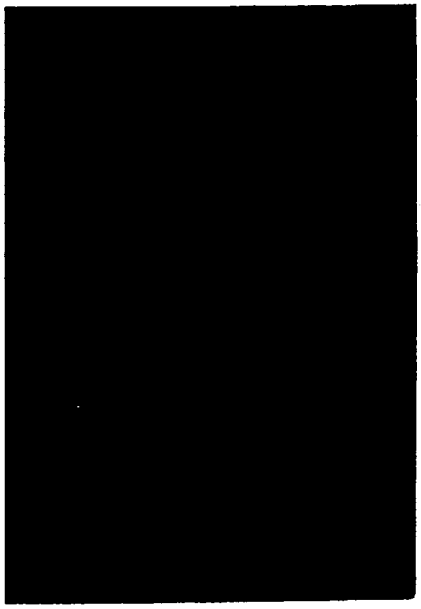


FIG.3B

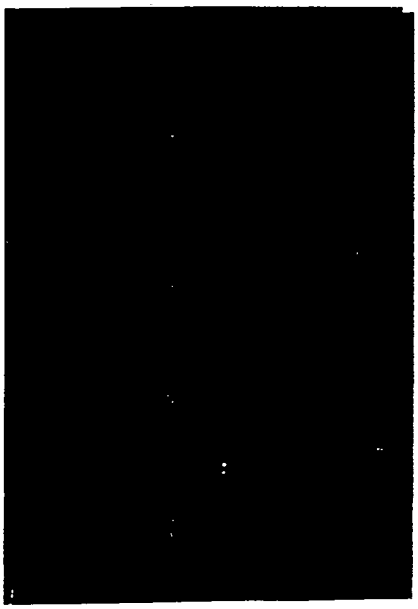


FIG.3C

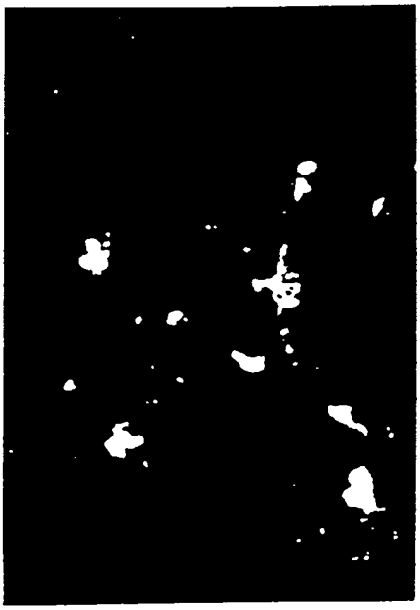


FIG.3D

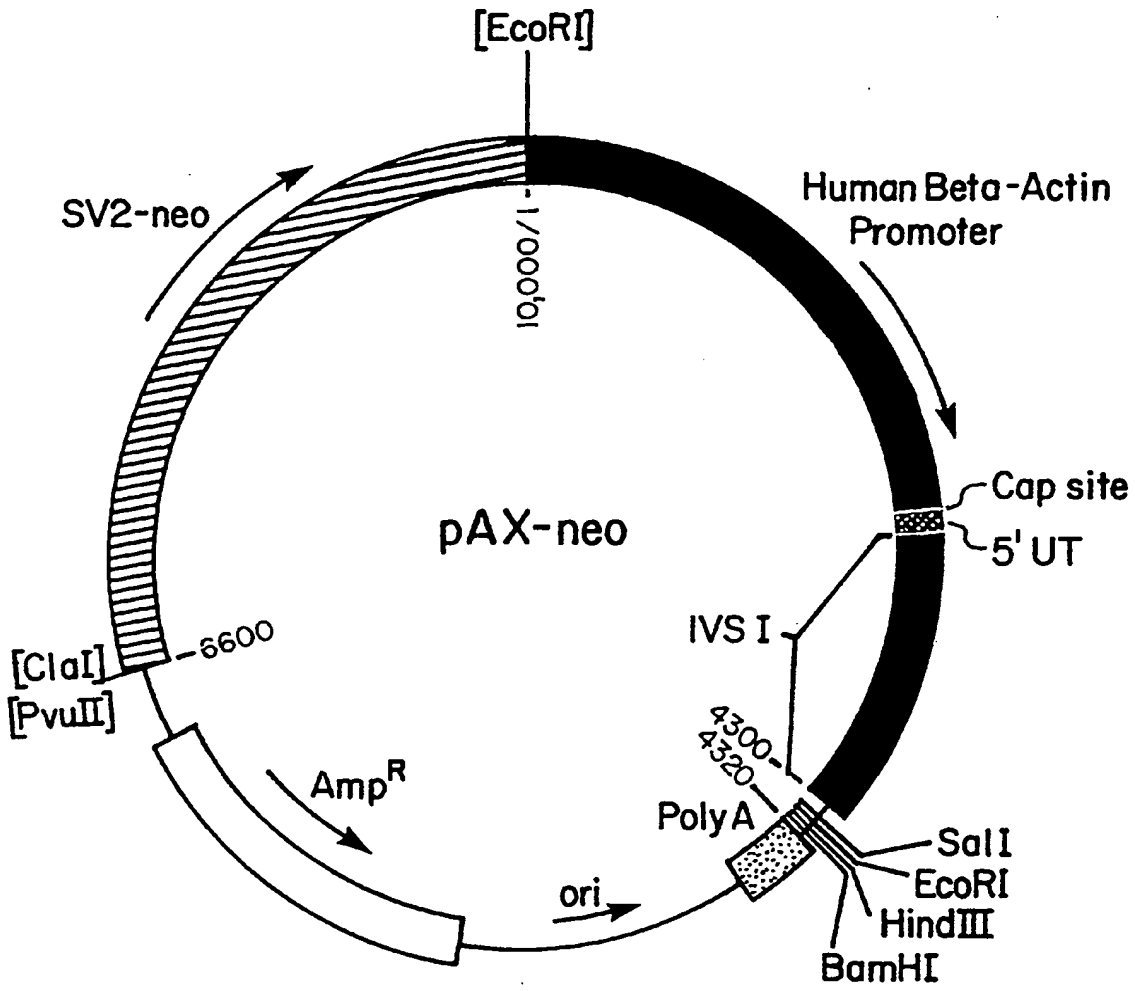

[EcoRI]
SV2-neo
Human Beta-Actin Promoter
1/000/01
pAX-neo
Cap site
5' UT
IVS I
[ClaI]
[PvuII]
6600
4300
4320
Amp^R
PolyA
SalI
EcoRI
HindIII
BamHI
ori

FIG. 4

INTERNATIONAL SEARCH REPORT

International Application No PCT/US90/05155

I. CLASSIFICATION OF SUBJECT MATTER (if several classification symbols apply, indicate all) [3]

According to International Patent Classification (IPC) or to both National Classification and IPC

IPC(5): C12Q 1/68; C12P 21/02; G01N 33/53, 33/531; A61K 35/14
US.CL: 435/6, 7, 70.1, 172.3, 810; 436/543, 548, 808; 530/350, 387, 839

II. FIELDS SEARCHED

Minimum Documentation Searched [4]

Classification System	Classification Symbols
U.S. CL.	435/6, 7, 70.1, 172.3, 810; 436/543, 548, 808; 530/350,387, 839

Documentation Searched other than Minimum Documentation to the Extent that such Documents are Included in the Fields Searched [5]

APS. CAS. BIOSIS

III. DOCUMENTS CONSIDERED TO BE RELEVANT [14]

Category *	Citation of Document, [16] with indication, where appropriate, of the relevant passages [17]	Relevant to Claim No. [18]
A, P	US,A, 4,919,915 (AVERBACK) 24 April 1990, See column 1, lines 15-22; column 1, line 59-column 2, line 32; column 2, lines 41-46 and 53-55 and column 9, lines 25-55.	1-14
A, P	US,A, 4,912,206 (GOLDGABER ET AL.) 27 March, 1990, see Figures 1 and 3 and claims 1-3.	1-6
X Y	US,A, 4,666,829 (GLENNER ET AL.) 19 May 1987, see Abstract and column 4, lines 5-11 and 19-29.	9-11 12-14
A	US,A, 4,264,729 (BELJANSKI) 28 April 1981, See Abstract and claim 1.	1-8
Y	WO,A, 89/07657 (NEVE ET AL.) 24 August 1989, see page 15, line 15-page 16, line 7.	9-14

* Special categories of cited documents: [15]

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

IV. CERTIFICATION

Date of the Actual Completion of the International Search [2]	Date of Mailing of this International Search Report [2]
26 November 1990	24 JAN 1991
International Searching Authority [1]	**Signature of Authorized Officer** [20]
ISA/US	Janelle Graeter Janelle Graeter

Form PCT/ISA/210 (second sheet) (May 1986)

III. DOCUMENTS CONSIDERED TO BE RELEVANT (CONTINUED FROM THE SECOND SHEET)

Category *	Citation of Document, [16] with indication, where appropriate, of the relevant passages [17]	Relevant to Claim No [18]
A Y	EP,A, 0,304,013 (KITAGUCHI ET AL) 22 February 1989, see Abstract, page 5, lines 43-44; page 7, line 45-page 8, line 5; page 8, line 30-page 9, line 12; page 11, line 47-page 12, line 31 and page 16, line 1-page 17, line 8.	1-8 9-14

From genes to protein structure and function: novel applications of computational approaches in the genomic era

Jeffrey Skolnick and Jacquelyn S. Fetrow

The genome-sequencing projects are providing a detailed 'parts list' of life. A key to comprehending this list is understanding the function of each gene and each protein at various levels. Sequence-based methods for function prediction are inadequate because of the multifunctional nature of proteins. However, just knowing the structure of the protein is also insufficient for prediction of multiple functional sites. Structural descriptors for protein functional sites are crucial for unlocking the secrets in both the sequence and structural-genomics projects.

Genome-sequencing projects are providing a detailed 'parts list' for life. Unfortunately, this list, a portion of which represents the amino acid sequence of all the proteins in a given genome, does not come with an instruction manual. That is, given the genome's sequences, one does not necessarily know straight away which regions encode proteins, which serve a regulatory role and which are responsible for the structure and replication of the DNA itself.

This is not unlike giving a child a list of parts necessary to create a working automobile. Without the necessary expertise, creating the final, working car from just the initial parts list is a nearly impossible task. Similarly, understanding how to create a complete, functioning cell given just the sequence of nucleotides found in an organism's genome is a complex problem.

What is a protein function?

After a genome is sequenced and its complete parts list determined, the next goal is to understand the function(s) of each part, including that of the proteins. What do we mean by protein function, the focus of this article?

Function has many meanings. At one level, the protein could be a globular protein, such as an enzyme, hormone or antibody, or it could be a structural or membrane-bound protein. Another level is its biochemical function, such as the chemical reaction and the substrate specificity of an enzyme. The regulatory molecules or cofactors that bind to a protein are also levels of biochemical function.

At the cellular level, the protein's function would involve its interaction with other macromolecules and the function and cellular location of such complexes. There is also the protein's physiological function; that is, in which metabolic pathway the protein is involved or what physiological role it performs in the organism. Finally, the phenotypic function is the role played by the protein in the total organism, which is observed by deleting or mutating the gene encoding the protein.

J. Skolnick (skolnick@danforthcenter.org) is at the Danforth Plant Science Center, Laboratory of Computational Genomics, 4041 Forest Park Avenue, St Louis, MO 63108, USA. J.S. Fetrow is at GeneFonnatics, Suite 200, 5830 Oberlin Drive, San Diego, CA 92121-3754, USA.

Obviously, the complete characterization of protein function is difficult but efforts are under way at all levels[1–4], including cellular function[5,6]. In this article, however, we focus on identifying the biochemical function of a protein given its sequence, a problem that is amenable to molecular approaches.

Sequence-based approaches to function prediction

The sequence-to-function approach is the most commonly used function-prediction method. This robust field is well developed and, in the interest of space limitations, we will merely present a brief overview.

There are two main flavors of this approach: sequence alignment[7–9]; and sequence-motif methods such as Prosite[10], Blocks[11], Prints[12,13] and Emotif[14]. Both the alignment and the motif methods are powerful but a recent analysis has demonstrated their significant limitations[15], suggesting that these methods will increasingly fail as the protein-sequence databases become more diverse.

An extension of these approaches that combines protein-sequence with structural information has been developed and some successes have been reported[16]. However, this method still applies the structural information in a one-dimensional, 'sequence-like' fashion and fails to take into account the powerful three-dimensional information displayed by protein structures.

In addition, proteins can gain and lose function during evolution and may, indeed, have multiple functions in the cell (Box 1). Sequence-to-function methods cannot specifically identify these complexities. Inaccurate use of sequence-to-function methods has led to significant function-annotation errors in the sequence databases[17].

An alternative approach

An alternative, complementary approach to protein-function prediction uses the sequence-to-structure-to-function paradigm. Here, the goal is to determine the structure of the protein of interest and then to identify the functionally important residues in that structure. Using the chemical structure itself to identify functional sites is more in line with how the protein actually works.

0167-7799/00/$ – see front matter PII: S0167-7799(99)01398-0

In a sense, this is one long-term goal of 'structural genomics' projects[18,19], which are designed to determine all possible protein folds experimentally, just as genome-sequencing projects are determining all protein sequences[20]. This is in contrast to traditional structural-biology approaches, in which one knows the protein's function first and only then, if the function is sufficiently important, determines its structure.

It is implicitly assumed that having the protein's structure will provide insights into its function, thereby furthering the goals of the human-genome-sequencing project. However, knowing a protein's three-dimensional structure is insufficient to determine its function (Box 2). What we really need to analyse and predict the multifunctional aspects of proteins is a method specifically to recognize active sites and binding regions in these protein structures.

Active-site identification

In order to use a structure-based approach to function prediction, one must identify the key residues responsible for a given biochemical activity. For many years, it has been suggested that the active sites in proteins are better conserved than the overall fold. Taken to the limit, this suggests that one could not only identify distant ancestors with the same global fold and the same activity but also proteins with similar functions but distantly related, or possibly unrelated, global folds.

The validity of this suggestion was demonstrated empirically by Nussinov and co-workers, who showed that the active sites of eukaryotic serine proteases, subtilisins and sulfhydryl proteases exhibit similar structural motifs[21]. Furthermore, in a recent modeling study of *Saccharomyces cerevisiae* proteins, protein functional sites were found to be more conserved than other parts of the protein models[22]. Similarly, it has been demonstrated that the catalytic triad of the α/β hydrolases is structurally better conserved than other histidine-containing triads[23]. A comparison of the structure of the hydrolase catalytic triad to other histidine-containing triads shows a distinct bimodal distribution, while a similar analysis done with a randomly selected triad shows a unimodal distribution (Fig. 1).

Kasuya and Thornton[24] generalized this example by creating structural analogs of a few Prosite sequence motifs[10]. For the 20 most-frequently occurring Prosite patterns, the associated local structure is quite distinct. These results provide clear evidence that enzyme active sites are indeed more highly conserved than other parts of the protein.

Identifying active sites in experimental structures

Historically, several groups have attempted to identify functional sites in proteins; these efforts were directed at protein engineering or building functional sites in places where they did not previously exist. This has been successfully accomplished for several metal-binding sites[25–33]. However, highly accurate functional-site descriptors of the backbone and side-chain atoms were required, fueling the belief that significant atomic detail is required in site descriptors for function identification.

Highly detailed residue side-chain descriptors of the active sites of serine proteases and related proteins have been used to identify functional sites[3]. The use of these highly detailed motifs has led to the identification of several novel functional sites in known, high-quality protein structures[3,34]. More automated methods for finding spatial motifs in protein structures have also been described[21,34–40].

Unfortunately, most of these methods require the exact placement of atoms within protein backbones and side chains, and so have not been shown to be relevant to inexact predicted structures. Recently, however, we described the production of fuzzy, inexact descriptors of protein functional sites[15]. As we wish to apply the descriptors to experimental structures as well as to predicted protein models, we used only carbon atoms and side-chain centers-of-mass positions. We call these descriptors 'fuzzy functional forms' (FFFs) and have created them for both the disulfide-oxidoreductase[15,41] and α/β-hydrolase catalytic active sites[23].

The disulfide-oxidoreductase FFF was applied to screen high-resolution structures from the Brookhaven protein database[42]. In a dataset of 364 protein structures, the FFF accurately identified all proteins known to exhibit the disulfide-oxidoreductase active site[15]. In a larger dataset of 1501 proteins, the FFF again accurately identified all proteins with the active site. In addition, it identified another protein, 1fjm, a serine–threonine phosphatase. This result was initially discouraging but subsequent sequence alignment and clustering analysis strongly suggested that this putative site might indeed be a site of redox regulation in the serine–threonine phosphatase-1 subfamily[43]. If confirmed by experiment, this result will highlight the advantages of using structural descriptors to analyse multiple functional sites in proteins. It will also highlight the fact that human-

Box 1. Proteins are multifunctional

A common protein characteristic that makes functional analysis based only on homology especially difficult is the tendency of proteins to be multifunctional. For instance, lactate dehydrogenase binds NAD, substrate and zinc, and performs a redox reaction. Each of these occurs at different functional sites that are in close proximity and the combination of all four sites creates the fully functional protein.

Other examples of multifunctional proteins are the nucleic-acid-binding proteins. For instance, DNA regulatory proteins often contain a DNA-binding domain, a multimerization domain and additional sites that bind regulatory proteins; a classic example is RecA[59]. The 3C rhinovirus protease exhibits a proteolytic function as well as an RNA-binding function[60,61]. Transcription factors are also complex, multifunctional proteins[62]. It is becoming increasingly important to recognize each of these different functions of gene products of a newly sequenced gene.

The serine–threonine-phosphatase superfamily is a prime example of the difficulties of using standard sequence analysis to recognize the multiple functions found in single proteins. This large protein family is divided into a number of subfamilies, all of which contain an essential phosphatase active site. Subfamilies 1, 2A and 2B exhibit 40% or more sequence identity between them[63]. However, each of these subfamilies is apparently regulated differently in the cell[64–67] and observation suggests that there are different functional sites at which regulation can occur. Because the sequence identity between subfamilies is so high, standard sequence-similarity methods could easily misclassify new sequences as members of the wrong subfamily if the functional sites are not carefully considered, as was recently demonstrated[43].

These are but a few examples of the multifunctionality of proteins. The recognition of this multifunctional nature is of critical importance to the genomics field. Useful functional-annotation methods must consider all of the specific functions in a given protein and will not just provide a general classification of function.

Box 2. Knowing a protein's structure does not necessarily tell you its function

Because proteins can have similar folds but different functions[68,69], determining the structure of a protein may or may not tell you something about its function. The most well-studied example is the $(\alpha/\beta)_8$ barrel enzymes, of which triose-phosphate isomerase (TIM) is the archetypal representative. Members of this family have similar overall structures but different functions, including different active sites, substrate specificities and cofactor requirements[70,71].

Is this example common? Our own analysis of the 1997 SCOP database[68] shows that the five largest fold families are the ferredoxin-like, the (α/β) barrels, the knottins, the immunoglobulin-like and the flavodoxin-like fold families with 22, 18, 13, 9 and 9 subfamilies, respectively (Fig. i). In fact, 57 of the SCOP fold families consist of multiple superfamilies. These data only show the tip of the iceberg, because each superfamily is further composed of protein families and each individual family can have radically different functions. For example, the ferredoxin-like superfamily contains families identified as Fe–S ferredoxins, ribosomal proteins, DNA-binding proteins and phosphatases, among others.

After this article was submitted, a much-more-detailed analysis of the SCOP database was published[72]. This finds a broad function–structure correlation for some structural classes, but also finds a number of ubiquitous functions and structures that occur across a number of families. The article provides a useful analysis of the confidence with which structure and function can be correlated[72]. Knowing the protein structure by itself is insufficient to annotate a number of functional classes and is also insufficient for annotating the specific details of protein function.

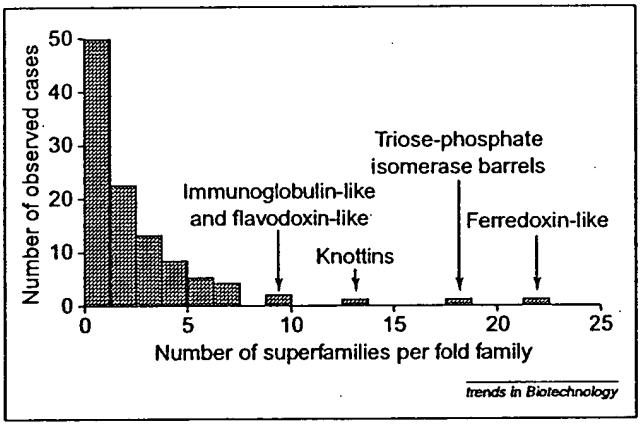


Figure I

Histogram of the numbers of superfamilies found in each SCOP fold family. These data clearly show that proteins with similar structures can have different functions and demonstrate the difficulty of assigning protein function based simply on the three-dimensional structure. The data were taken from the 1997 distribution of SCOP (http://scop.mrc-lmb.cam.ac.uk/scop). For a more-detailed analysis, see Ref. 72.

observation alone is no longer adequate for identifying all functional sites in known protein structures.

To date, the use of structure to identify function has largely focused on high-resolution structures and highly detailed descriptors of protein functional sites. However, the creation of inexact descriptors for functional sites opens the way to the application of these methods to inexact, predicted protein models. The question remains: how good does a model have to be in order to use FFFs to identify its active sites?

The state of the art in structure-prediction methods

For proteins whose sequence identity is above ~30%, one can use homology modeling to build the structure[44]. However, structure prediction is far more difficult for proteins that are not homologous to proteins with known structure. At present, there are two approaches for these sequences: *ab initio* folding[45–48] and threading[49–53].

In *ab initio* folding, one starts from a random conformation and then attempts to assemble the native structure. As this method does not rely on a library of pre-existing folds, it can be used to predict novel folds. The recent CASP3 protein-structure-prediction experiment (http://PredictionCenter.llnl.gov/CASP3) involved the blind prediction of the structure of proteins whose actual structure was about to be experimentally determined. These results indicate that considerable progress has been made[46,54]. For helical and α/β proteins with less than 110 residues, structures were often predicted whose backbone root-mean-square deviation (RMSD) from native ranged from 4–7 Å. Progress is being made with the β proteins, too, although they remain problematic. Because *ab initio* methods can identify novel folds, these methods could be used to help to select sequences likely to yield novel folds in experimental structural-genomics projects.

Another approach to tertiary-structure prediction is threading. Here, for the sequence of interest, one attempts to find the closest matching structure in a library of known folds[52,55]. Threading is applicable to proteins of up to 500 residues or so and is much faster than *ab initio* approaches. However, threading cannot be used to obtain novel folds.

Ab initio *predicted models can be used for automatic protein-function prediction*

The results of the recent CASP3 competition suggest that current modeling methods can often (but not always) create inexact protein models. Are these structures useful for identifying functional sites in proteins? Using the *ab initio* structure-prediction program MONSSTER, the tertiary structure of a glutaredoxin, 1ego, was predicted[56]. For the lowest-energy model, the overall backbone RMSD from the crystal structure was 5.7 Å.

To determine whether this inexact model could be used for function identification, the sets of correctly and incorrectly folded structures were screened with the FFF for disulfide-oxidoreductase activity[15]. The FFF uniquely identified the active site in the correctly folded structure but not in the incorrectly folded ones (Fig. 2). This is a proof-of-principle demonstration that inexact models produced by *ab initio* prediction of structure from sequence can be used for the subsequent prediction of biochemical function. Of course, improvements in the method have to be made before such predictions can be done on a routine basis.

Use of predicted structures from threading in protein-function prediction

At present, practical limitations preclude folding an entire genome of proteins using *ab initio* methods[57]. Threading is more appropriate for achieving the requisite high-throughput structure prediction. Thus, a standard threading algorithm[58] has been used to screen all

proteins in nine genomes for the disulfide-oxidoreductase active site described above.

First, sequences that aligned with the structures of known disulfide oxidoreductases were identified. Then, the structure was searched for matches to the active-site residues and geometry. For those sequences for which other homologs were available, a sequence-conservation profile was constructed[23]. If the putative active-site residues were not conserved in the sequence subfamily to which the protein belongs, that sequence was eliminated. Otherwise, the sequence is predicted to have the function.

Using this sequence-to-structure-to-function method, 99% of the proteins in the nine genomes that have known disulfide-oxidoreductase activity have been found. From 10% to 30% more functional predictions are made than by alternative sequence-based approaches; similar results are seen for the α/β hydrolases[23]. Surprisingly, in spite of the fact that threading algorithms have problems generating good sequence-to-structure alignments, active sites are often accurately aligned, even for very distant matches. This observation would agree with the above experimental results indicating that active sites are well conserved in protein structures.

Importantly, the false-positive rate when using structural information is much lower than that found using sequence-based approaches, as demonstrated by a detailed comparison of the FFF structural approach and the Blocks sequence-motif approach (N. Siew *et al.*, unpublished). In this study, the sequences in eight genomes, including *Bacillus subtilis*, were analysed for disulfide-oxidoreductase function using the disulfide-oxidoreductase FFF, the thioredoxin Block 00194 and the glutaredoxin Block 00195. If we assume that those sequences identified by both the FFF and Blocks are 'true positives', we find 13 such sequences in the *B. subtilis* genome.

There is no experimental evidence validating all of these 'true positives' and so they are more accurately termed 'consensus positives'. In order to find these 13 'consensus positive' sequences, the FFF hits seven false positives. On the other hand, Blocks hits 23 false positives (Fig. 3). It was previously suggested that the use of a functional requirement adds information to threading and reduces the number of false positives[52]. These data, including the data shown in Fig. 3, validate this claim on a genome-wide basis.

Of course, as no genome has had the function of all of its proteins experimentally annotated, it is impossible to know how many other proteins with the specified biochemical function were not properly identified. This is a critical question for researchers attempting to predict protein function. Experimental confirmation will be needed to validate this or any other method fully. This points out the need for closely coupling computational function-prediction algorithms with experiments.

Weaknesses of using the sequence-to-structure-to-function method of function prediction

Based on studies to date, the identification of enzymatic activity requires a model in which the backbone RMSD from native near the active sites is about 4–5 Å. Predicted models are better at describing the geometry in the core of the molecule than in the loops and so predicting the function of a protein whose active site is in loops may be a problem. Also, the method can currently only be applied to enzyme active sites; substrate- and ligand-binding sites have not been identified using the inexact models. Techniques that will further refine inexact protein models will be quite useful in taking the protein analysis to the next step.

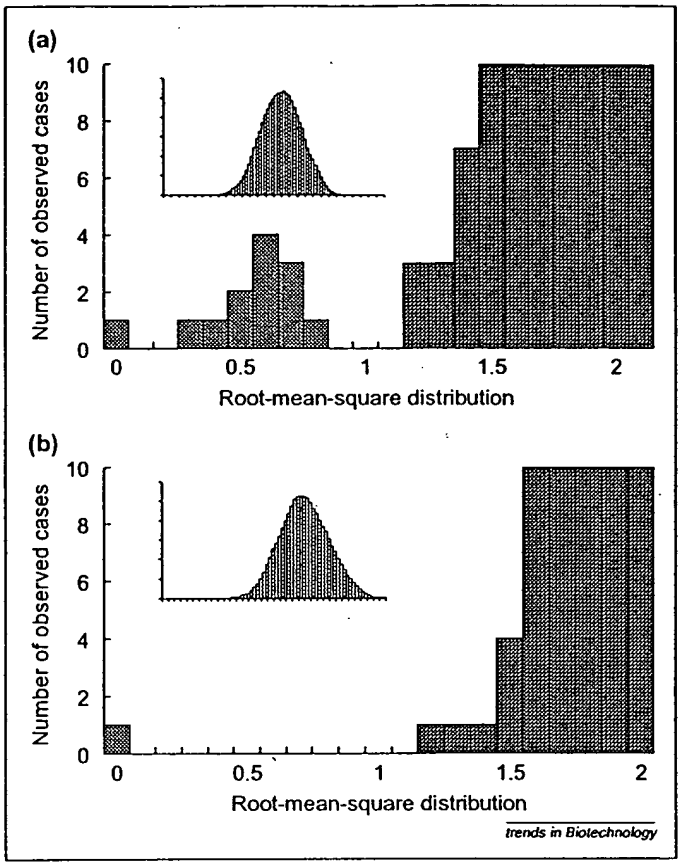


Figure 1

The distribution of root-mean-square distributions (RMSD) between the hydrolase catalytic triad and all other histidine-containing triads shows a bimodal distribution **(a)**; by contrast, the RMSD between a randomly selected (non-catalytic) triad and all other histidine-containing triads has a unimodal distribution **(b)**. The His–Ser–Asp catalytic trial in the protein-1 gpl (Rp2 lipase) **(a)** and a random histidine-containing triad from 4pga (glutaminase–asparaginase) **(b)** were structurally aligned to all His-containing triads in a database of 1037 proteins[23]. Actual α/β-hydrolase active sites **(a)** and the 4pga site **(b)** are indicated by blue bars; other histidine triads that are not active sites are indicated by red bars. None of the sites found by matching to the 4pga were hydrolase active sites. Inset graphs show the full distribution.

Conclusions

Although sequence-based approaches to protein-function prediction have proved to be very useful, alternatives are needed to assign the biochemical function of the 30–50% of proteins whose function cannot be assigned by any current methods. One emerging approach involves the sequence-to-structure-to-function paradigm. Such structures might be provided by structural-genomics projects or by structure-prediction algorithms. Functional assignment is made by screening the resulting structure against a library of structural descriptors for known active sites or binding regions.

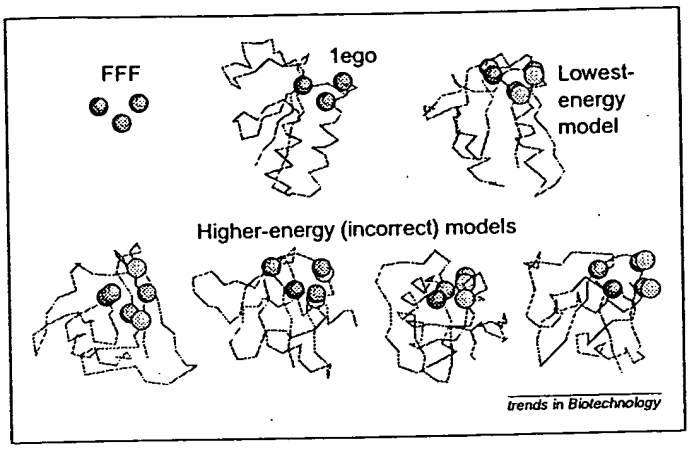


Figure 2

Application of the disulfide-oxidoreductase fuzzy functional form (FFF) to *ab initio* models of glutaredoxin created by the program MONSSTER shows that the FFF can distinguish between correctly folded and misfolded (or higher-energy) models. The FFF is shown as two orange balls (representing the cysteines) and a blue ball (representing the proline). The protein models are shown as magenta wire models with the active-site cysteines and proline shown as yellow and cyan balls, respectively. The FFF clearly distinguishes the correct active site in the crystal structure of the glutaredoxin 1ego and the correctly folded, lowest-energy model. The FFF does not match to the active sites of any of the higher energy, misfolded structures, four of which are shown here.

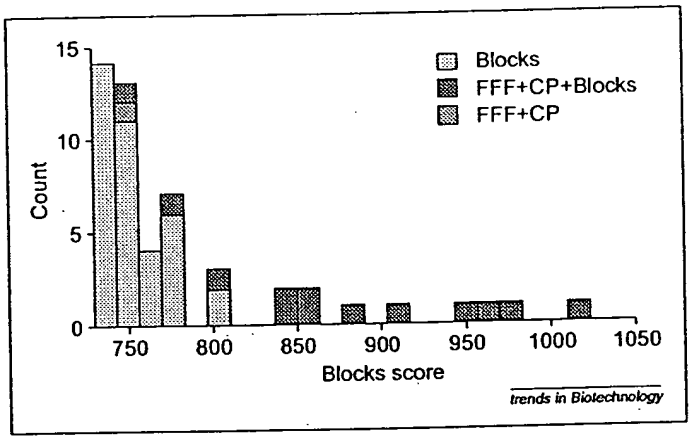


Figure 3

Analysis of the *Bacillus subtilis* genome using the thioredoxin Block 00194. The Blocks score (computed using the publicly available BLIMPS program) is plotted on the x axis and the number of sequences found in each scoring bin is plotted on the y axis. Those sequences identified as 'consensus positives' [identified by both the fuzzy functional form (FFF) and the Block] are shown as red bars. One additional sequence found by the FFF, which is likely to be a true positive, is shown as a blue bar. All other sequences, putative 'false positives', are shown as yellow bars. Using the Blocks score at which all 13 of the 'consensus positives' are found, 23 false positives are also found. In its analysis of the *B. subtilis* genome, the FFF identifies only seven false positives along with the same 13 'consensus positives' (data not shown).

Detailed descriptors will only work on the experimentally determined, high-quality structures. Ideally, however, the descriptors should work on both experimental structures and the cruder models provided by tertiary-structure-prediction algorithms.

The advantages of such an approach are that one need not establish an evolutionary relationship in order to assign function, that more than one function can be assigned to a given protein [an issue of major importance, because proteins are multifunctional (Box 1)] and, ultimately, that having a structure can provide deeper insight into the biological mechanism of protein function and regulation. The disadvantages are that one needs to have the protein's structure before a function can be assigned and that the approach is limited to those functions associated with proteins with at least one solved structure, so that a functional-site descriptor can be constructed.

In this sense, structure-to-function assignment can be thought of as 'functional threading' – find the active-site match in a library of descriptors for known protein active sites. This is the first step in the long process of using structure to assign all levels of function, a goal that is made increasingly important with the emergence of structural genomics. Based on the progress to date, it is apparent that structure will play an important role in the post-genomic era of biology.

Acknowledgment

We thank L. Zhang for producing the data in Box 2 and Fig. 1.

References

1 Gurd, F.R.N. and Rothgeb, T.M. (1979) Motions in proteins. *Adv. Protein Chem.* 33, 73–165
2 Laskowski, R.A. *et al.* (1996) X-SITE: use of empirically derived atomic packing preferences to identify favourable interaction regions in the binding sites of proteins. *J. Mol. Biol.* 259, 175–201
3 Wallace, A.C. *et al.* (1996) Derivation of 3D coordinate templates for searching structural databases: application to Ser–His–Asp catalytic triads in the serine proteinases and lipases. *Protein Sci.* 5, 1001–1013
4 Henikoff, S. and Henikoff, J.G. (1991) Automated assembly of protein blocks for database searching. *Nucleic Acids Res.* 19, 6565–6572
5 Riley, M. (1993) Functions of gene products of *Escherichia coli*. *Microbiol. Rev.* 57, 862–952
6 Karp, P.D. and Riley, M. (1993) Representations of metabolic knowledge. *Ismb* 1, 207–215
7 Altschul, S.F. *et al.* (1990) Basic local alignment search tool. *J. Mol. Biol.* 215, 403–410
8 Pearson, W.R. (1996) Effective protein sequence comparison. *Methods Enzymol.* 266, 227–258
9 Sturrock, S.S. and Collins, J.F. (1993) *Biocomputing Research Unit*, University of Edinburgh, Edinburgh, UK
10 Bairoch, A. *et al.* (1995) The PROSITE database, its status in 1995. *Nucleic Acids Res.* 24, 189–196
11 Henikoff, S. and Henikoff, J.G. (1994) Protein family classification based on searching a database of blocks. *Genomics* 19, 97–107
12 Attwood, T.K. *et al.* (1994) PRINTS – A database of protein motif fingerprints. *Nucleic Acids Res.* 22, 3590–3596
13 Attwood, T.K. *et al.* (1997) Novel developments with the PRINTS protein fingerprint database. *Nucleic Acids Res.* 25, 212–216
14 Nevill-Manning, C.G. *et al.* (1998) Highly specific protein sequence motifs for genome analysis. *Proc. Natl. Acad. Sci. U. S. A.* 95, 5865–5871
15 Fetrow, J.S. and Skolnick, J. (1998) Method for prediction of protein function from sequence using the sequence-to-structure-to-function paradigm with application to glutaredoxins/thioredoxins and T1 ribonucleases. *J. Mol. Biol.* 281, 949–968
16 Yu, L. *et al.* (1998) A homology identification method that combines protein sequence and structure information. *Protein Sci.* 7, 2499–2510
17 Bork, P. and Bairoch, A. (1996) Go hunting in sequence databases but watch out for traps. *Trends Genet.* 12, 425–427
18 Gaasterland, T. (1998) Structural genomics: bioinformatics in the driver's seat. *Nat. Biotechnol.* 16, 625–627
19 McKusick, V.A. (1997) Genomics: structural and functional studies of genomes. *Genomics* 45, 244–249
20 Montclione, G.T. and Anderson, S. (1999) Structural genomics: keystone for a human proteome project. *Nat. Struct. Biol.* 6, 11–12

21 Fischer, D. *et al.* (1994) Three-dimensional, sequence order-independent structural comparison of a serine protease against the crystallographic database reveals active site similarities: potential implications to evolution and to protein folding. *Protein Sci.* 3, 769–778

22 Sanchez, R. and Sali, A. (1998) Large-scale protein structure modeling of the *Saccharomyces cerevisiae* genome. *Proc. Natl. Acad. Sci. U. S. A.* 95, 13597–13602

23 Zhang, L. *et al.* (1998) Functional analysis of *E. coli* proteins for members of the α/β hydrolase family. *Fold. Design* 3, 535–548

24 Kasuya, A. and Thornton, J.M. (1999) Three-dimensional structure analysis of Prosite patterns. *J. Mol. Biol.* 286, 1673–1691

25 Coldren, C.D. *et al.* (1997) The rational design and construction of a cuboidal iron–sulfur protein. *Proc. Natl. Acad. Sci. U. S. A.* 94, 6635–6640

26 Pinto, A.L. *et al.* (1997) Construction of a catalytically active iron superoxide dismutase by rational protein design. *Proc. Natl. Acad. Sci. U. S. A.* 94, 5562–5567

27 Hellinga, H.W. and Richards, F.M. (1991) Construction of new ligand binding sites in proteins of known structure: (I) computer-aided modeling of sites with pre-defined geometry. *J. Mol. Biol.* 222, 763–785

28 Hellinga, H.W. *et al.* (1991) Construction of new ligand binding sites in proteins of known structure: (II) grafting of a buried transition metal binding site into *Escherichia coli* thioredoxin. *J. Mol. Biol.* 222, 787–803

29 Klemba, M. and Regan, L. (1995) Characterization of metal binding by a designed protein: single ligand substitutions at a tetrahedral Cys_2His_2 site. *Biochemistry* 34, 10094–10100

30 Klemba, M. *et al.* (1995) Novel metal-binding proteins by design. *Nat. Struct. Biol.* 2, 368–373

31 Farinas, E. and Regan, L. (1998) The *de novo* design of a rubredoxin-like Fe site. *Protein Sci.* 7, 1939–1946

32 Crowder, M.W. *et al.* (1995) Spectroscopic studies on the designed metal-binding sites of the 43C9 single chain antibody. *J. Am. Chem. Soc.* 117, 5627–5634

33 Halfon, S. and Craik, C.S. (1996) Regulation of proteolytic activity by engineered tridentate metal binding loops. *J. Am. Chem. Soc.* 118, 1227–1228

34 Wallace, A.C. *et al.* (1997) TESS: A geometric hashing algorithm for deriving 3D coordinate templates for searching structural databases: application to enzyme active sites. *Protein Sci.* 6, 2308–2323

35 Kleywegt, G.J. (1999) Recognition of spatial motifs in protein structures. *J. Mol. Biol.* 285, 1887–1897

36 Matsuo, Y. and Nishikawa, K. (1994) Protein structural similarities predicted by a sequence-structure compatibility method. *Protein Sci.* 3, 2055–2063

37 Russell, R.B. (1998) Detection of protein three-dimensional side-chain patterns: new examples of convergent evolution. *J. Mol. Biol.* 279, 1211–1227

38 Han, K.F. *et al.* (1997) Three-dimensional structures and contexts associated with recurrent amino acid sequence patterns. *Protein Sci.* 6, 1587–1590

39 Artymiuk, P.J. *et al.* (1994) A graph-theoretic approach to the identification of three-dimensional patterns of amino acid side-chains in protein structures. *J. Mol. Biol.* 236, 327–344

40 Karlin, S. and Zhu, Z.Y. (1996) Characterizations of diverse residue clusters in protein three-dimensional structures. *Proc. Natl. Acad. Sci. U. S. A.* 93, 8344–8349

41 Fetrow, J.S. *et al.* (1998) Functional analysis of the *Escherichia coli* genome using the sequence-to-structure-to-function paradigm: identification of proteins exhibiting the glutaredoxin/thioredoxin disulfide oxidoreductase activity. *J. Mol. Biol.* 282, 703–711

42 Abola, E.E. *et al.* (1987) *Protein Data Bank in Crystallographic Databases: Information Content, Software Systems, Scientific Application* (Allen, F.H. *et al.*, eds), Data Commission of the International Union of Crystallography, Bonn/Cambridge/Chester

43 Fetrow, J.S. *et al.* (1999) Structure-based functional motif identifies a potential disulfide oxidoreductase active site in the serine/threonine protein phosphatase-1 subfamily. *FASEB J.* 13, 1866–1874

44 Sali, A. *et al.* (1995) Evaluation of comparative protein modeling by MODELLER. *Proteins* 23, 318–326

45 Bystroff, C. and Baker, D. (1998) Prediction of local structure in proteins using a library of sequence-structure motifs. *J. Mol. Biol.* 281, 565–577

46 Shortle, D. (1999) The state of the art. *Curr. Biol.* 9, R205–R209

47 Lee, J. *et al.* (1999) Calculation of protein conformation by global optimization of a potential energy function. *Proteins* 3 (Suppl.), 204–208

48 Ortiz, A. *et al.* (1999) *Ab initio* folding of proteins using restraints derived from evolutionary information. *Proteins* 3 (Suppl.), 177–185

49 Bowie, J.U. *et al.* (1991) A method to identify protein sequences that fold into a known three-dimensional structure. *Science* 253, 164–170

50 Finkelstein, A.V. and Reva, B.A. (1991) A search for the most stable folds of protein chains. *Nature* 351, 497–499

51 Bryant, S.H. and Lawrence, C.E. (1993) An empirical energy function for threading protein sequence through folding motif. *Proteins* 16, 92–112

52 Lathrop, R. and Smith, T.F. (1996) Global optimum protein threading with gapped alignment and empirical pair scoring function. *J. Mol. Biol.* 255, 641–665

53 Jones, D.T. *et al* (1992). A new approach to protein fold recognition. *Nature* 358, 86–89

54 Sternberg, M.J. *et al.* (1999) Progress in protein structure prediction: assessment of CASP3. *Curr. Opin. Struct. Biol.* 9, 368–373

55 Miller, R.T. *et al.* (1996) Protein fold recognition by sequence threading tools and assessment techniques. *FASEB J.* 10, 171–178

56 Ortiz, A.R. *et al.* (1998) Fold assembly of small proteins using Monte Carlo simulations driven by restraints derived from multiple sequence alignments. *J. Mol. Biol.* 277, 419–448

57 Skolnick, J. *et al.* (1998) Reduced protein models and their application to the protein folding problem. *J. Biomol. Struct. Dyn.* 16, 381–396

58 Jaroszewski, L. *et al.* (1998) Fold prediction by a hierarchy of sequence, threading and modeling methods. *Protein Sci.* 7, 1431–1440

59 Takahashi, M. *et al.* (1996) Locations of functional domains in the RecA protein: overlap of domains and regulation of activities. *Eur. J. Biochem.* 242, 20–28

60 Leong, L.E. *et al.* (1993) Human rhinovirus-14 protease 3C (3Cpro) binds specifically to the 5′ noncoding region of the viral RNA: evidence that 3Cpro has different domains for the RNA binding and proteolytic activities. *J. Biol. Chem.* 268, 25735–25739

61 Matthews, D.A. *et al.* (1994) Structure of human rhinovirus 3C protease reveals a trypsin-like polypeptide fold, RNA-binding site and means for cleaving precursor polyprotein. *Cell* 77, 761–771

62 Ladomery, M. (1997) Multifunctional proteins suggest connections between transcriptional and post-transcriptional processes. *BioEssays* 19, 903–909

63 Goldberg, J. *et al.* (1995) Three-dimensional structure of the catalytic subunit of protein serine/threonine phosphatase-1. *Nature* 376, 745–753

64 Mumby, M.C. and Walter, G. (1993) Protein serine/threonine phosphatases: structure, regulation and functions in cell growth. *Physiol. Rev.* 73, 673–699

65 Jia, Z. (1997) Protein phosphatases: structures and implications. *Biochem. Cell Biol.* 75, 17–26

66 Holmes, C.F.B. and Boland, M.P. (1993) Inhibitors of protein phosphatase-1 and -2A: two of the major serine/threonine protein phosphatases involved in cellular regulation. *Curr. Opin. Struct. Biol.* 3, 934–943

67 Nemani, R. and Lee, E.Y.C. (1993) Reactivity of sulfhydryl groups of the catalytic subunits of rabbit skeletal muscle protein phosphatases 1 and 2A. *Arch. Biochem. Biophys.* 300, 24–29

68 Murzin, A.G. *et al.* (1995) Scop: a structural classification of proteins database for the investigation of sequences and structures. *J. Mol. Biol.* 247, 536–540

69 Orengo, C.A. *et al.* (1997) CATH: a hierarchic classification of protein domain structures. *Structure* 5, 1093–1108

70 Lesk, A.M. *et al.* (1989) Structural principles of α/β proteins: the packing of the interior of the sheet. *Proteins Struct. Funct. Genet.* 5, 139–148

71 Farber, G.K. and Petsko, G.A. (1990) The evolution of α/β barrel enzymes. *Trends Biochem. Sci.* 15, 228–234

72 Hegyi, H. and Gerstein, M. (1999) The relationship between protein structure and function: a comprehensive survey with application to the yeast genome. *J. Mol. Biol.* 288, 147–164